MUSICAL ABSTRACTIONS FOR MULTI-ROBOT COORDINATION

A Thesis
Presented to
The Academic Faculty

by

María Teresa Santos Fernández

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in
Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
May 2016

MUSICAL ABSTRACTIONS FOR MULTI-ROBOT COORDINATION

Approved by:

Dr. Magnus B. Egerstedt, Advisor
Schlumberger Professor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. George F. Riley
Professor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Ayanna M. Howard
Professor and Linda J. and Mark C. Smith Chair in Bioengineering
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: April 25, 2016

ACKNOWLEDGMENTS

I want to thank my advisor, Dr. Magnus Egerstedt, for giving me the amazing opportunity of working on this project and for having faith in its potential. Without his enthusiastic encouragement I would have never believed my music theory inquisitiveness could be incorporated into this robotics research journey.

I also want to thank the members of the GRITS Lab for their warm welcome, their support with this unconventional research and their invaluable help in setting up the experimental work. This last year has been truly a pleasure.

Finally, I want to acknowledge the Fulbright Commission in Spain and the Iberdrola Foundation for their support of my career aspirations. I feel very grateful for having the privilege of being chosen to embark on this amazing academic experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This work presents a new approach to human-swarm interactions, a discipline which addresses the problem of how a human operator can influence the behavior of large groups of robots, providing high-level information understandable by the team. While there exist potential advantages of introducing a human in the control loop of a robot swarm, how the human must be incorporated is not a simple problem. For the intervention of a human operator to be favorable to the performance of the team, the means and form of the information between the human and the robot swarm must be adequately defined: we need to design which device will be provided to the operator to interact with the swarm and how the information will be shaped so that both the human and the robot team understand it.

Coordination of multi-robot systems involves the generation of involved motion patterns for the individual agents that result in an overall organized movement. We introduce in this thesis a new human-swarm interaction modality based on music theory, a discipline studied for centuries and capable of creating complex sound structures. In particular, we have focused on understanding how we can apply rules and structures from music theory to an operator's input so that each command both specifies the goal location to be visited and the geometry to be adopted by the swarm. We interpret the sequence of locations to be visited by the swarm as a musical melody, identifying each note with a certain location in the robots' workspace. Once the objective path is defined in the form of a melody, we can apply rules from harmony, a discipline of music theory, to create chords that harmonize the input melody. The interest in using these chords lies fundamentally in that they are structured combinations of pitches, heard simultaneously. These inherent structures will be used to determine the geometry that should be displayed by the team. The developed multi-robot control is applied to a team of differential drive mobile robots through an electronic piano.

CHAPTER 1
INTRODUCTION

As computing becomes smaller, faster and cheaper, autonomous robotic systems have also reduced their size and augmented their capabilities. As a result, multi-robot coordination has flourished as a new approach for complex applications such as environmental monitoring and surveillance, due to its inherent robustness and scalability. In applications where a large number of robots collaborate, the capabilities of the team are often enhanced with the introduction of a human operator, capable of managing and injecting high-level information and allowing the system to rapidly adapt to unplanned situations. However, while there exist situations where the inclusion of a person can have a beneficial outcome for the multi-robot application, the question of how this information should be provided to the team remains unanswered.

This thesis introduces a new approach to the problem of human-swarm interaction in which we apply music theory as the abstraction to model the information flow between the human user and the robot team. In particular, our objective is to apply broadly studied composition techniques to create, from simple operator inputs, complex commands capable of coordinating the movement of a multi-robot team. In this chapter we introduce the topic of human-swarm interaction, which constitutes the main focus of this thesis due to its relevance in multi-agent systems. We also include a detailed description of the problem we aim to overcome with this work, as well as an outline of the organization of this thesis.

1.1 The Problem of Human-Swarm Interaction

As multi-robot systems are being envisioned and deployed to handle tasks of increasing complexity and scope, such as environmental monitoring and surveillance [1, 2], intelligent warehousing [3], or collaborative manufacturing and materials handling [4], human involvement with the robot team is becoming increasingly paramount. Although a number of basic building blocks for controlling teams of mobile robots in a collaborative fashion have been designed, e.g., rendezvous [5, 6], area coverage [7, 8], and formation control [9, 10, 11], the topic of human-swarm interactions is still lacking a unified treatment, e.g. [12].

Human-swarm interactions focus on the question of how teams of mobile robots should be organized in order to make themselves amenable to effective human influence. This includes not only the robots' interaction laws, but also the search for effectively manipulable "abstractions", such as leader-follower networks [13, 14], virtual leaders and other abstract points of control [15], fluid-based swarm abstractions [16], or behavioral abstractions [9, 17]. In conjunction with this, the question must be addressed of how these abstractions should be manipulated, i.e., what types of "swarm joysticks" should be used. In the literature, a number of such interaction tools have been proposed, ranging from classical joysticks and computer interfaces [18, 19, 17], through haptic interfaces [20, 21], to deformable clay [22].

The choice of interaction modality depends on factors such as the size of the swarm (do individuals matter?) and the type of tasks under consideration. In this work, we have investigated a novel interaction modality that allows a user to dictate *where* the swarm should go and *what* its spatial arrangement should look like. To this end, we use a "tool" that has already been extensively studied and explored for achieving complex effects, namely music theory. In particular, we center our attention on harmony rules as a way to modify the control parameters of the swarm and illustrate the use of these musical interaction abstractions by controlling a team of mobile robots using a piano. This idea is illustrated in Figure 1 where we can observe a human user controlling the movement of a robot team in its workspace through the piano keyboard, which serves as the interaction device.

Figure 1: Implementation of the human-swarm interaction modality presented in this work. The human user introduces high-level commands through a piano keyboard. The user's input is interpreted by the application program, which generates the motion commands that make the team approach an area of interest while displaying a formation (a triangle in this case).

1.2 Problem Description

The goal of this thesis is to explore the viability of a human-swarm interaction modality that uses music theory as an abstraction for multi-robot control. The importance of human-swarm interaction for the development of applications of swarm robotics, along with the nonexistence of a unified solution on the topic were the main motivation behind the development of this work. The well documented study of music theory during the last four centuries along with its ability to create complex musical effects make it a good candidate to be explored and introduced in the field.

The specific problem we want to solve is the following: given a team of robots, we want to give the human operator the ability of specifying a sequence of target positions the team must visit. In addition, while advancing towards the specified goals, we do not want the group to herd but, instead, to proceed in an orderly manner, adopting a geometric formation. However, in order to avoid a task overload for the operator, we want to restrict their responsibilities to the solely specification of the areas of interest. The required isolation of the user from the specification of the formation geometries is the motivation for the use of music theory. Music theory allows us to create sound combinations from the specification of single melody lines. In other words, it gives us tools to create complex structures (the geometries) from simple inputs. Thus, if we interpret the sequence of positions as notes in a melody line, harmony rules can be applied to such melody to obtain harmonizing chords. We can then use the structure of those chords to define the geometry of the formations to be adopted.

The design of this human-swarm interaction methodology involves the completion of four main subtasks. First, we need to perform a deep study of the musical concepts involved in the harmonization process, as it will allow us to use harmony as the abstraction to model the human input as control commands understandable by the robotic platform. Second, we want the system to generate the suitable harmonies for the operator's input. For that purpose, the applicable harmony rules need to be mathematically modeled in such a way that they become implementable in a computer algorithm. Third, a map between the musical structures generated by the algorithm and the actual robot movement must be designed. Finally, the performance of the developed interaction methodology needs to be tested on a real robotic platform, in order to analyze the adequacy of music theory as an abstraction for multi-robot control.

1.3 Organization of the Document

This thesis documents the chosen approach to develop the the proposed human-swarm interaction strategy. This section has been included to provide an outline of the document, where the content of each chapter in this document has been summarized.

First, a literature review of the related work to this thesis is provided in Chapter 2, where a summary of the main contributions in the field of human-swarm interactions has been outlined along with a description of different works where music and multi-agent systems have been combined. Chapter 3 provides an introduction to graph theoretic methods for multi-agent systems, which will be applied in later chapters to control the movement of the members of the team.

To fulfill the first subtask of the thesis, we provide in Chapter 4 an outline of the main music theory concepts used in the course of this work. The description of the developed mathematical model based on automata, responsible for the automatic harmonization of the user's input, is also included in this chapter. The mapping criteria between the harmonies and the formation geometry to be displayed by the robot team is the scope of Chapter 5. In addition to the mapping, the formation control laws are also introduced in this chapter to completely overcome the gap between the musical structures and the actual robot movement to satisfy the team's formation specification.

Chapter 6 presents the experimental work carried out in this thesis, where the mathematical model introduced in Chapter 4 is combined with the mapping between the musical abstractions and the formation control described in Chapter 5, resulting in a human-swarm interaction that has been tested on two real robotic platforms.

The conclusions of this thesis are detailed in Chapter 7, where we summarize the main contributions of this work. To conclude this document we have included Chapter 8, where the main lines through which this work could be continued are presented.

CHAPTER 2
PREVIOUS WORK

During the last decades, an increasing interest towards swarm robotics has arisen. Swarm's inherent reliability and distributed character make them the preferred candidate in applications such as rendezvous [5, 6], area coverage [7, 8], and formation control [9, 10, 11]; where their performance excel against single high-equipped robots. However, while the swarm of robots is intended to autonomously carry out their assigned tasks, environmental difficulties are subject to arise in these applications and prevent the swarm from fulfilling its objectives. Such situations can be overcome by introducing a human operator, able to provide with high-level information.

The attractive of having a swarm of robots working cooperatively with a human lies on the potential adaptability of goals for the swarm as well as on the overcome of possible hazards or difficulties that could prevent the swarm from completing its tasks. However, in order to take advantage of these benefits, the means of the interaction need to be defined. Human-swarm interaction investigates how a person's input can be combined with the control laws of the swarm in order to improve the performance of the swarm and achieve a desirable output.

This work introduces a new modality of human-swarm interaction based on music theory. The aim of this chapter is to describe the related work existent in the literature as of today. A summary of the different approaches in the field of human-swarm interaction is presented together with a description of some works that have combined robotic swarms and music.

2.1 Human-Swarm Interaction

How one single operator should influence the performance of a large group of robots is not a simple question in that the answer depends on a number of different factors, ranging from the composition of the team [23] to the complexity of the task [12]. Making a large group of robots carry out a complex and dynamic task involves considerable amounts of decision making that often exceed the cognitive and reasoning capabilities of a human operator, e.g. [24]. For the sake of an efficient interaction of the operator with the robot team, the design of the interaction must isolate the operator from managing the individual actions of the robots [18]. Instead, the role of the operator is to deal with higher-level interactions that enhance the performance of the swarm such as defining areas of interest, changing objectives or protecting the swarm from potential hazards [7, 18, 24].

Human-swarm interactions have received significant attention during the last decade. Some approaches propose for the human operator to interact with the swarm by identifying the human with one of the robots in the swarm, e.g. [25]. The commanded robot by the operator can be treated equally among the other agents or have a higher category than them, affecting their behavior through attraction and repulsion effects [17], for example. In behavior-based swarm control, the interactions are inspired by communication mechanisms observed in certain animal colonies[26] allows human operator to position artificial beacons, acting like animal pheromones. Body gestures have been also explored as an interaction mechanism for the swarm, e.g.[27].

As for the means of interaction with the selected abstraction, different instruments and materials have been explored as potential joysticks. Computer interfaces constitute a common approach due to the familiarity of the user with this kind of devices, together with the capabilities of providing visual feedback [25, 8, 19]. On the other hand, deformable materials like clay have been explored due to their flexibility in terms of being manually modeled by the human operator, adopting different shapes and being splitted or merged in various pieces of material [22]. The effectiveness of the operator commands is often another concern on the topic of human-swarm interaction, which can be addressed considering providing feedback to the operator through haptic devices [20].

The aforementioned interaction approaches provide the user with capabilities for directing the swarm towards areas of interest, or avoid hazardous areas, in addition to producing other group behaviors such as splitting and merging. However, these approaches do not address any dynamic control of the structure of the robot team. In this work, we propose a new approach for human-swarm interaction that uses music theory to generate movement and formations for a group of robots. We take advantage of the sound distance structure present in chords and apply it to robot formation. The human operator interfaces with the robotic system by playing a note in a piano keyboard, from which the chord structure will be generated according to music theory rules. A leader-follower strategy is chosen to illustrate the music-based control, where the piano input specifies target locations to be visited and the formation geometry is selected according to the corresponding chord.

2.2 Music and Robots in the Literature

Music has been applied to the robotics world in numerous ocassions, with anthropomorphic robots playing musical instruments being the best known example of this combination of disciplines. Vaucanson's Flute Player, an automaton presented in 1738 at the Académie des Sciences, constitutes the first attempt to create a robotic musician [28]. Following this precedent, various anthropomorphic robots capable of playing musical instruments have been created over the last decades. We can find examples of anthropomorphic playing robots in each of the three instrument families. In 1974, a robot capable of playing a concert organ was presented [29]. Within the winds we can also find flutist [30], saxophonist [31, 32], trumpetist and trombonist robots [31]. Violin often constitutes the most popular approach when creating humanoid robot musicians [33, 34, 35], although some cello examples have also been designed [35]. Percussions instruments have also been explored [36] since they allow for improvisational musical applications as opposed to the previous examples, where the robotic platforms often play pieces that consist in sequences of preprogrammed movements with respect to the musical instrument. Less traditional instruments like the electronic theremin have also been played by humanoid robots [37]. In general, with instrument playing robots, the adaptability of the platforms is usually conditioned by their mechanical complexity, which results in a limited interchangeability of instruments [31].

Apart from mimicking the human motion that arises from playing a musical instrument, other aspects like musical improvisation between humans and robots have been addressed. The marimba playing robot Shimon [38, 39, 40] goes a step further in musical interpretation. As a response to a human-produced musical sequence, Shimon is capable of analyzing the human's style of playing to which it responds with a jazz-styled improvisation that follows within the style standards of the human's input. Even in this case the robot was given a humanoid appearance, since studies reflect that it seems to increase the satisfaction of the audience [36], as opposed to purely computer synthesized sounds. Other examples involve human and robot musical interaction, for example by interpreting musical rhythms [41] or as a practice and compositional aid for human musicians [42].

Orchestral robotic examples can also be found in the literature. One of the most significant examples is the Man and Machine robot orchestra at the Logos Foundation in Ghent, Belgium [43]. This orchestra, composed of more than 40 automated instruments have served as a resource for experimental composers and musicians since 1968. The experimental and innovative character of robotic orchestras is also illustrated in Pat Meheny's orchestrion [44], a group of acoustic and electroacoustic instruments that are automatically controlled and that allows for jazz improvisations.

2.2.1 Music and Swarm Robotics

The combination of music and robots, mainly in the form of instrument playing robots, has been extensively explored for the last four decades. However, the topic of human-swarm interaction has received limited attention from a musical perspective.

When blending music with swarm robotics, one option is to consider music as the output of a multi-robot system that emerges from the interaction of the agents with the environment [45] or between themselves [46]. In [45], music is generated by the combination of preprogrammed sounds that each robot reproduces when completing a particular task. However, the sounds to be played by the robots need not be preprogrammed. A music generation algorithm was implemented in each of the robots in [46], so that different rhythm patterns were played depending on the interaction of each robot with its team neighbors. Improvisational music systems based on the concurrent deployment of various swarms has also been explored [47].

Music can be used to characterize the role of the different robots within a swarm [48]. When a polyphonic musical piece is interpreted by the robotic swarm by playing predefined sound tracks, agents can be clustered into subgroups depending on their role within the musical piece. Music has been also used as an inspiration for dealing with temporal-constrained routing of groups of robots [23], where robots need to hit sets of locations at specific times, as a musician has to play the different notes in a musical score at the appropriate times.

In this work, the combination of music and swarm robotics has a different character. Instead of using the interactions between robots to generate music, music theory serves a generator of formation control commands for the robot team. In particular, we apply well established harmony structures to the input specified by the human operator. From this process we obtain structured combination of sounds that we translate into geometries to be displayed by the robot team.

CHAPTER 3

MULTI-AGENT ROBOTICS

Coordinating the movement of a multi-robot team involves the definition of the local interaction rules that lead to the desired global behavior. In the proposed human-swarm interaction approach we need to define the control protocols to ensure the team accomplishes the leader-follower formation control strategy in order to illustrate the music-based control.

A graph-based approach is used in this work to generate the control laws applicable to formation control. In this chapter we provide a condensed description of the theoretic methods related to the leader-follower formation control implemented in this work. We begin by introducing the main concepts of graph theory used in networked control. Next, we introduce rendezvous, a canonical example that illustrates the links between the abstractions in graph theory and the actual physics of the distributed system. Finally, we will adapt the control laws designed for the rendezvous problem, shaping them to be applied in our desired formation control strategy.

3.1 Graph-Based Network Models

Multi-robot systems are characterized by the presence of multiple agents that interact with each other in a network to achieve system-level objectives [49]. Agents in those networks do not usually possess a general picture of the status of the network. Instead, they gather information from local interactions with other agents, by sensing or communicating with them. In order to model this interactions we use graphs, a mathematical tool that provides natural abstractions for the information flow between agents in a network. In this section we introduce the main elements of graph theory that are involved in the definition of the control laws for our robot team, providing intuition about how to map the mathematical abstractions into the characteristics of the physical system.

3.1.1 What is a Graph?

A *graph* is a representation of a finite set of elements, which constitute the *vertex set*, denoted by V, where each element of V is a *vertex*. When the vertex set has n elements, we can represent it as

$$V = \{v_1, v_2, \dots, v_n\} \tag{1}$$

The vertices in a graph are related in pairs. Consider the set of pairwise elements of V, denoted by $[V]^2$, which consists of elements of the form $\{v_i, v_j\}$ such that $i, j = 1, 2, \dots, n$ and $i \neq j$. The finite graph $\mathcal{G}$ is defined as

$$\mathcal{G} = (V, E) \tag{2}$$

where E is a particular subset of $[V]^2$, called the set of *edges* of $\mathcal{G}$, which represents the connections between the elements in V. Therefore, a graph can mathematically encode the structure of a multi-robot system. In our case, each of the vertices of the graph represents a robot in the team and two vertices will share an edge if the two represented robots interact in the network.

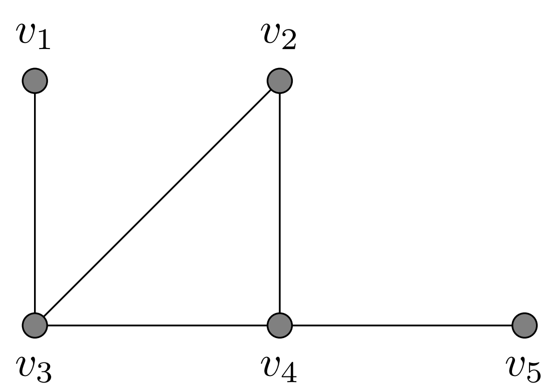


Figure 2: An undirected graph with 5 vertices.

A graph is a set theoretic object. However, as its name suggests, it allows for graphical representation. The graphical representation of $\mathcal{G}$ consists of dots representing the vertices of the graph and lines between the vertices when there exists an edge that connects them.

Figure 2 gives an example of an undirected graph, $\mathcal{G} = (V, E)$, where $V = \{v_1, v_2, \ldots, v_5\}$ and $E = \{v_1v_3, v_2v_3, v_2v_4, v_3v_4, v_4v_5\}$. If this graph was the representation of our multi-robot system, then our system would be formed by five robots represented by V. The edge set E provides information about the relationships in the network: robot 1 only interacts with robot 3; robot 2 with robots 3 and 4; robot 3 with robots 1, 2 and 4; robot 4 with robots 2, 3 and 5; and robot 5 only with robot 4.

The graphical representation of a graph leads to many definitions and observations about graphs including adjacency, incidence and connectedness. If two vertices, v_i and v_j, share an edge they are called *adjacent*, which is denoted by $v_i \sim v_j$. Reciprocally, in such case, the edge v_iv_j is *incident* with vertices v_i and v_j. Introducing the concept of adjacency allows for the definition of neighborhoods for the vertex set. The *neighborhood* $N(i) \subseteq V$ of the vertex v_i is understood as the set $\{v_j \in V \mid v_iv_j \in E\}$, that is, the set of vertices that are adjacent to v_i. If $v_j \in N(i)$, it follows that $v_i \in N(j)$, since the edge set in an undirected graph consists of unordered vertex pairs. For example, in the graph example shown in Figure 2, v_2 and v_3 are adjacent, and the neighborhood of v_2 is $N(2) = v_3, v_4$.

We call a graph $\mathcal{G}$ *connected* if there is a path along the edges between any two vertices. The graph in Figure 2 is connected since we can go from any vertex in the graph to another by following the defined edges. On the other hand, Figure 3 shows an example of a disconnected graph where it is impossible to go from v_2 to v_5, for example.

3.1.2 Algebraic Graph Theory

We have seen in the previous section how graphs allow for representing relationships between a finite number of objects, while admitting a straightforward graphical representation in terms of vertices and edges. One of the main advantages of graphs is that they also allow representations in terms of matrices, providing the possibility of manipulating the mathematical abstractions of the networked system through matrix algebraic operations.

The concepts of adjacency and neighborhoods, as well as connectedness, can be represented and analyzed through matrices. For an undirected graph $\mathcal{G} = (V, E)$, the cardinality

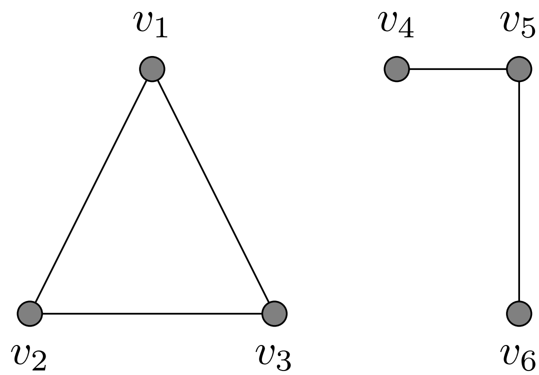


Figure 3: Example of a disconnected graph with 6 vertices.

of the neighborhood set of a given vertex v_i is referred to as *degree*

$$d(v_i) = |N(i)| \tag{3}$$

For the graph shown in Figure 2 the degrees of the vertices are: $d(v_1) = 1, d(v_2) = 2, d(v_3) = 3, d(v_4) = 3, d(v_5) = 1$. The *degree matrix* of a graph $\mathcal{G}$ with n vertices is the diagonal matrix containing the degrees of its vertices, $d(v_i)$, on the diagonal, that is,

$$\Delta(\mathcal{G}) = \begin{pmatrix} d(v_1) & 0 & \dots & 0 \\ 0 & d(v_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d(v_n) \end{pmatrix} \tag{4}$$

In our robotic application, if we assign one identifiers to each of the robots, we can encode or check the number of robots interacting with a particular robot i by simply looking at the entry $[\Delta(\mathcal{G})]_{ii}$ of the degree matrix.

On the other hand, while the degree matrix indicates the number of neighbors for each agent in the network, the *adjacency matrix* encodes the identity of those neighbors. The adjacency matrix $A(\mathcal{G})$ is the symmetric $n \times n$ matrix encoding the adjacency relationships in the graph $\mathcal{G}$, in that

$$[A(\mathcal{G})]_{ij} = \begin{cases} 1 \text{ if } v_i v_j \in E \\ \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

Therefore, for a robot i in the multi-robot system, we can know who are its neighbors by checking the entries of the row i in the adjacency matrix.

If we return to the example shown in Figure 2, the corresponding degree and adjacency matrices are

$$\Delta(\mathcal{G}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad A(\mathcal{G}) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{6}$$

The information contained in both the degree and adjacency matrices is included in the *graph Laplacian*, $L(\mathcal{G})$, a matrix which provides all the information about the structure of a graph. Although there exist other ways to obtain the Laplacian for a graph (see [49]), we provide in this work the most straightforward specification of the graph Laplacian associated with an undirected graph $\mathcal{G}$. We can define the graph Laplacian as

$$L(\mathcal{G}) = \Delta(\mathcal{G}) - A(\mathcal{G}) \tag{7}$$

where $\Delta(\mathcal{G})$ and $A(\mathcal{G})$ are the degree and the adjacency matrix, respectively. From this definition, it trivially follows that all the rows of a graph Laplacian sum to zero, for any graph. The importance of the graph Laplacian lies in its spectral information, since we can infer various properties of the network by analyzing its eigenvalues. For example, we can study the connectedness of a graph by observing the number of zero eigenvalues of L, being the number of connected groups in the graph equal to the algebraic multiplicity of the zero as an eigenvalue. The graph Laplacians for the graphs in Figures 2 and 3 are

$$L(\mathcal{G}_2) = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} \qquad L(\mathcal{G}_3) = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \tag{8}$$

If we calculate their eigenvalues, we will observe that $L(\mathcal{G}_2)$ has one zero eigenvalue, corresponding to the connected component that represents the whole graph; while $L(\mathcal{G}_3)$ has two zero eigenvalues corresponding to the two connected components that can be observed in Figure 3: $\{v_1, v_2, v_3\}$ and $\{v_4, v_5, v_6\}$.

Therefore, by using graph representations and their associated matrix we can analyze different properties of the structure of a multi-robot system. These matrices provide a powerful tool in terms of analyzing the performance of different control laws implemented in the network.

3.2 Rendezvous: a Canonical Problem

Agreement is one of the fundamental problems in multi-agent coordination. Also referred to as the consensus or rendezvous problem, the objective of the agreement protocol is to get the agents to agree on a joint state value. In this section we describe the main insights of the rendezvous problem in order to provide a solid ground upon which we will define the formation control laws in Section 3.3.

The agreement protocol represents a canonical problem in the study of multi-agent systems in that it has a tight relation with many multi-agent problems such as flocking, swarming and distributed estimation. Assume a team of n dynamic agents which interact according to an undirected connected graph. If we denote the scalar state of unit i as $x_i \in \mathbb{R}$, we can achieve consensus by defining the rate of change of unit i as the sum of its relative states with respect to its neighboring units, that is

$$\dot{x}_i(t) = \sum_{j \in N(i)} (x_j(t) - x_i(t)), \quad i = 1, \ldots, n \tag{9}$$

where $N(i)$ is the set of unit i's neighbors in the network. Since we are assuming undirected graphs, the notion of adjacency is symmetric. Thus, the overall system can be represented by

$$\dot{x}(t) = -L(\mathcal{G})x(t) \tag{10}$$

where the positive semidefinite matrix $L(\mathcal{G})$ is the graph Laplacian of the agents interaction network (G) and $x(t) = (x_1(t), \ldots, x_n(t))^T \in \mathbb{R}$. The connectedness of the network plays a very important role in this problem since the agreement protocol presented at a node and ensemble level in Eq. 9 and 10 achieves consensus, that is,

$$\lim_{t \to \infty} (x_i(t) - x_j(t)) = 0 \quad \forall i, j \tag{11}$$

if and only if $\mathcal{G}$ is connected.

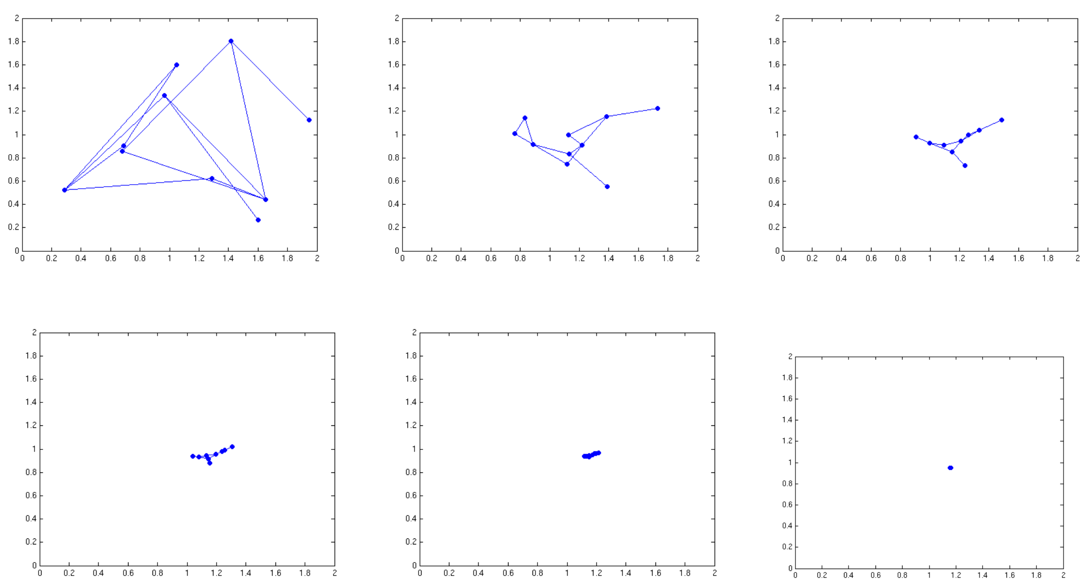


Figure 4: Simulation of the rendezvous protocol in a network of 10 planar agents. The agents start at random positions in the plane and interact according to an undirected graph, approaching the meeting point according to the consensus protocol.

The same idea can be translated to a team of planar robots which interact according to an undirected connected graph. In this case, the state of each robot i is its position, $x_i \in \mathbb{R}^2$. Assuming the dynamics of each robot in the team correspond to the single integrator dynamics, we can directly control the position through

$$\dot{x}_i = u \tag{12}$$

Then we can apply Eq. 9 to each of the two dimensions of the state. The results from applying this equation to a team of 10 agents whose structure corresponds to an undirected connected graph is presented in Figure 4.

3.2.1 Rendezvous Through Edge Energies

The canonical rendezvous presented in Eq. 9 can be manipulated to obtain more interesting behaviors than consensus such as ensuring the connectivity maintenance in the network or maintaining minimum distances between the agents. In this section we introduce changes on the consensus equation through the incorporation of weights, which will modify the behavior of the network.

A *weighted graph*, $\mathcal{G} = (V, E, w)$, results from adding to the vertex and edge sets a function $w : E \to \mathbb{R}$ that associates a value to each edge. Figure 5 depicts a weighted undirected graph with five nodes. Note that, since the graph is undirected, $w_{ij} = w_{ji}, \forall v_i v_j \in E$. Under this definition, the node-level dynamics of the consensus equation now becomes

$$\dot{x}_i(t) = \sum_{j \in N(i)} w_{ij}(x_j(t) - x_i(t)), \quad i = 1, \ldots, n \tag{13}$$

which will achieve consensus (see Eq. 11) if $\mathcal{G}$ is connected and $w_{ij} = w_{ji} > 0$.

But, how do we choose the weights in a network? The weights in Eq. 13 often adapt dynamically to the objectives of the multi-robot system, and thus become a function of time: $w_{ij}(t) = w_{ji}(t)$. In order to pick weight values in a standardized way, we can associate an energy $\mathcal{E}_{ij}(x_i, x_j) \in \mathbb{R}$ to each edge $v_i v_j$ in the graph, such that they are positive, $\mathcal{E}_{ij} > 0 \ \forall i \neq j$; and symmetric, $\mathcal{E}_{ij} = \mathcal{E}_{ji}$. Under this definition, the total energy in the network

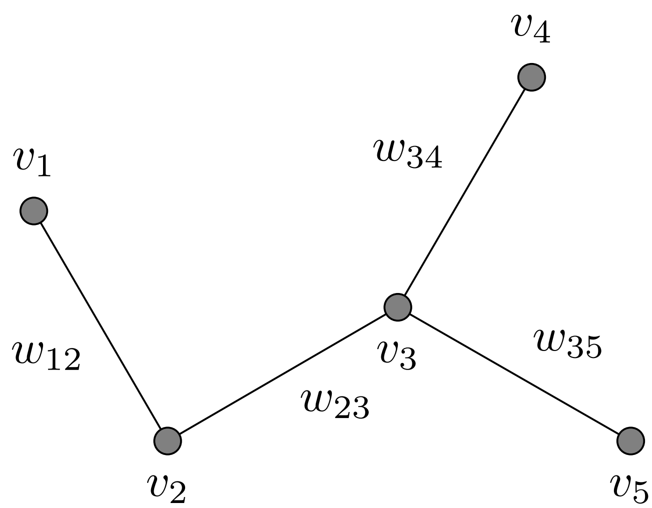


Figure 5: An undirected weighted graph with 5 vertices.

can be defined as

$$\mathcal{E} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \in N_i} \mathcal{E}_{ij}(x_i, x_j) \tag{14}$$

The edge energies $\mathcal{E}_{ij}$ are defined so that the global objective of the multi-robot system is attained when the total energy of the system is minimized. Therefore, we will want each agent to choose their actions aiming to decrease the total energy $\mathcal{E}$. This condition can be mathematically fulfilled by picking the dynamics of each agent as the negative gradient flow of the total energy:

$$\dot{x}_i = -\frac{\partial \mathcal{E}}{\partial x_i} \tag{15}$$

Deriving this expression according to the definition of total energy in Eq. 14, we find that

$$\dot{x}_i = -\sum_{j \in N(i)} \frac{\partial \mathcal{E}_{ij}}{\partial x_i} \tag{16}$$

which means that the overall objective is attained when each agent changes its state according to the negative gradient flow of the energies in its neighborhood. Picking the energy as a function of the distance between agents, that is, $\mathcal{E}_{ij}(x_i, x_j) = \mathcal{E}_{ij}(\|x_i - x_j\|)$ we get a weighted consensus equation

$$\dot{x}_i = \sum_{j \in N(i)} w_{ij}(\|x_i - x_j\|)(x_j - x_i) \tag{17}$$

To give a specific example, a possible application of this energy-driven consensus protocol could be ensuring the connectivity of the network when performing consensus. In a multi-robot system with a large number of agents, it is often the case that the robots have a limited sensor footprint, Δ, beyond which they cannot detect the presence of other robots. While performing a rendezvous protocol, it is vital that none of the agents loses contact with the rest of the network, since that would result in the system not achieving consensus. Thus, in such case the maintenance of all the edges in the graph at all times should be prioritized, while the global objective still remains to be meeting in a common point. One way of fulfilling these conditions is to define the edge energies so that they attain a minimum when the distance between the robots is zero and become very large when the connectivity with their neighbors is about to be lost. Mathematically, this could be written as

$$\mathcal{E}_{ij}(\|x_i - x_j\|) = \frac{\|x_i - x_j\|^2}{\Delta - \|x_i - x_j\|} \tag{18}$$

Figure 6 depicts a graphical representation of the edge energy defined in Eq. 18. We can see how, in effect, the value of the energy becomes very large when the distance between the agents approaches the value of the sensor footprint and becomes zero when they meet.

3.3 Formation Control

In the previous section we have laid the groundwork for the specification of the formation control laws, by analyzing the consensus protocol and introducing weights and edge energies. Similarly to the procedures described in Section 3.2.1, we will establish how to define edge energies in order to make the robot team display a formation geometry.

A formation geometry is achieved by a team of planar robots when they are located satisfying a set of geometry constraints. Figure 7 depicts an example of a formation geometry for a team of 8 robots where $d_{ij} = 1$ for all $v_i v_j \in E$ except for d_{13}, which is $\sqrt{2}$. In general,

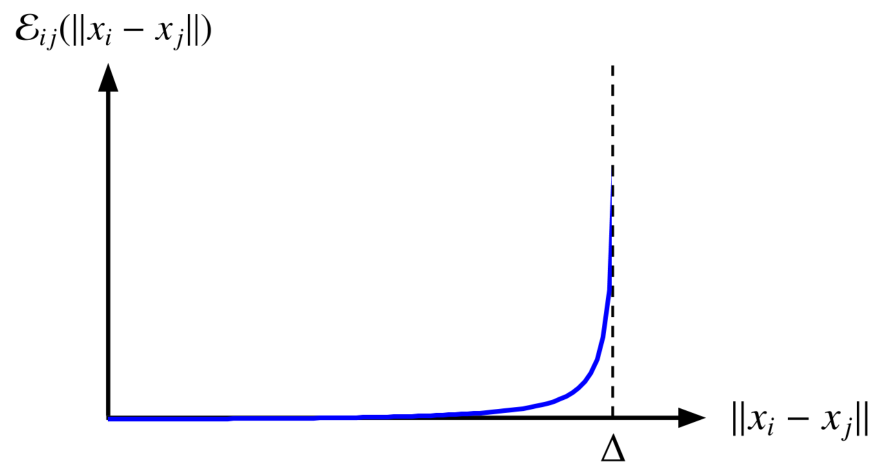


Figure 6: Edge energy profile for a consensus protocol with connectivity maintenance.

for a team of n planar robots, we can define a formation shape by establishing a set D of pairwise distances d_{ij} that the robots must maintain when displaying such formation. In the planar case, D is feasible if $\exists\ p_i, i = 1, \ldots, n,\ p_i \in \mathbb{R}^2$ such that $\|p_i - p_j\| = d_{ij}$, which means that the required distances between the agents must be consistent with a planar geometric figure. For example, we could not specify $d_{13} = 1$ in the formation of Figure 7, since it would be geometrically impossible for a planar formation to satisfy all the other distances specified in D.

We can define the *formation graph* as $\mathcal{G}_F = (V, E, w_{ij}(d_{ij}))$, where each of the edge weights, w_{ij}, depends on the desired distance between agents, d_{ij}. For the robot team to achieve a desired planar geometry, the formation graph needs to be rigid. By definition, a graph is *infinitesimally rigid* if the pairwise distances between agents are constant over time, that is

$$\|x_i(t) - x_j(t)\| = \|x_i(\tau) - x_j(\tau)\| \quad \forall t, \tau \in \mathbb{R}; \quad \forall v_i, v_j \in V \tag{19}$$

The rigidity conditions for a graph depend on the dimension in which the system lays. In the planar case we can ensure a graph is rigid if its structure is a planar triangulation, that is, it only consists of triangles without holes. The shape displayed in Figure 7 is rigid since it is a triangulation without holes. However, if we eliminate, for example, the

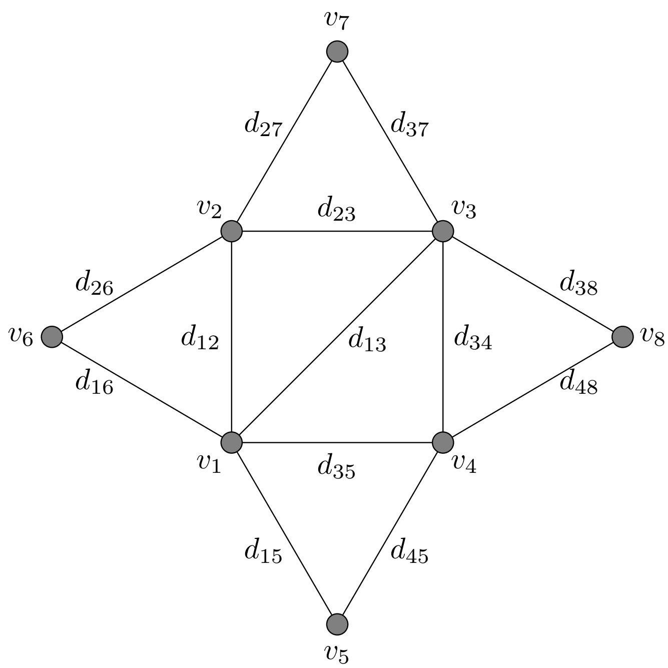


Figure 7: Example of a rigid formation specification by pairwise distances.

distance requirement d_{13} we create a hole and the graph becomes non-rigid. We illustrate this situation in Figure 8, where the graph has become non-rigid and different geometric configurations of the agents still satisfy the required pairwise distances.

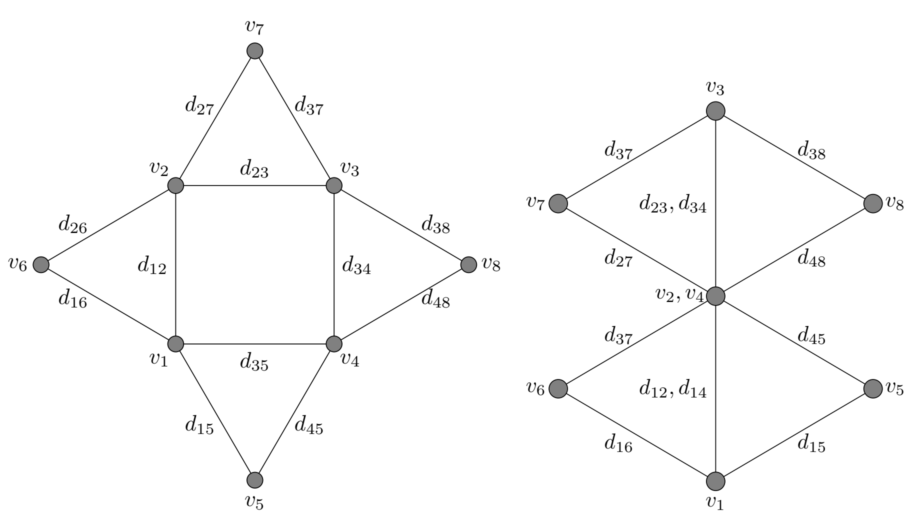


Figure 8: Two space configurations of the same non-rigid formation specification. Removing the diagonal of the central square results in a loss of rigidity: all the other distances in the structure are maintained even when the square is deformed into a rhombus.

In order to make the robots in the team display certain shape, we can modify the rendezvous consensus protocol in Eq. 17 by redefining the edge energies that yield the weights. In this case, the edge energies are designed so that the minimum of total energy, $\mathcal{E}$, is attained when all the agent systems are at their corresponding predefined distances. If we define the edge energy as

$$\mathcal{E}_{ij} = \frac{1}{2}\left(\|x_i - x_j\| - d_{ij}\right)^2 \tag{20}$$

then the resulting formation control equation is

$$\dot{x}_i = \sum_{\substack{j=1,\ldots N \\ j \neq i}} w_{ij}(\|x_i - x_j\|)(x_j - x_i) \tag{21}$$

with

$$w_{ij}(\|x_i - x_j\|) = \frac{\|x_i - x_j\| - d_{ij}}{\|x_i - x_j\|} \tag{22}$$

The resulting energy profile of the definition in Eq. 20 is shown in Figure 9. In this case, the minimum of the edge energy $\mathcal{E}_{ij}$ is attained when the relative distance between agents i and j is the pairwise distance of the formation specification, d_{ij}

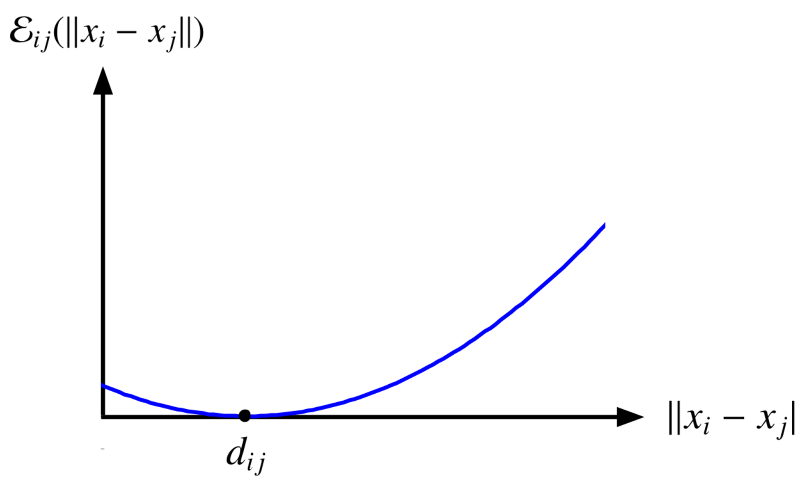


Figure 9: Edge energy profile for the formation control protocol.

We will use the formation control defined in Eq. 21 in Chapter 5, where the harmony structures introduced in this work will be mapped into robot movement.

CHAPTER 4

ROBOTS AND MUSIC THEORY

Music theory provides patterns, rules and even formulae for how to combine rhythms and harmonies to create the harmonious sequences that comprise what we call music. However, music is a complex art and a musical piece typically blends together multiple elements such as pitch, dynamics, timbre, consonance, texture, rhythm, etc. For the sake of complexity management, we will focus on only two of the aforementioned elements for controlling teams of mobile robots, namely the *consonance/dissonance* qualities of certain intervals[1] along with the *pitch* of the sounds. With these two elements we will be able to explore the main concepts of harmony, discipline that deals with the simultaneous combination of pitches, named *chords*, and their evolution towards other pitches over time. In this work, we will apply the principles of harmony used during the Common Practice Period of Western classical music (approx. 1600-1925), which comprises the Baroque, Classical, Romantic and Impressionst eras [50].

The human-swarm interaction strategy presented in this thesis is depicted in Figure 10. The human operator is given a piano keyboard as a ”joystick“, with which the operator can produce a melody for controlling the robot team. The input of the human directly controls one of the robots, selected as the leader, directing it to goal positions in the physical world. The remaining robots act as followers and depict a formation around the leader. The geometry of the formation depends on the human's input: for each note in the melody, the system will obtain the chord that harmonizes it best and use the structural properties of that chord to generate the geometry of the formation. The power of music theory resides on the variety of possibilities when harmonizing a melody, since we can obtain very different harmonizations using the same musical-elements palette. There does not exist a one-to-one relationship between a note in the melody and a chord, which implies that there is not a

[1] An *interval* represents the distance between two pitches. The minimum unit in Western classical music is a semitone, being a tone composed by two semitones.

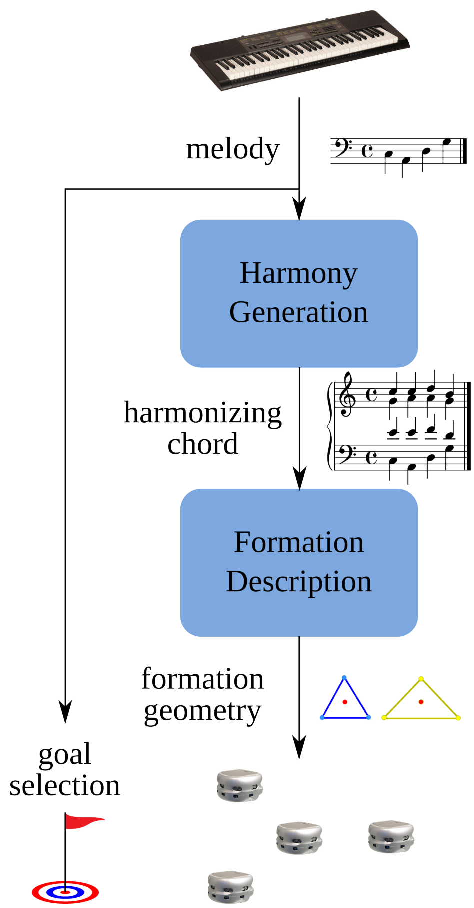


Figure 10: Transformation process of the operator's input melody into formation geometries for the robot team. The note played in the piano keyboard also indicates the target location to be visited.

unique formation geometry specification per each input of the operator.

This chapter focuses on the first part of the transformation of the operator's input, the generation of harmonizing chords from the user's input. We will first introduce the music theory concepts necessary for understanding the harmonizing process applied in this work. In the second half of this chapter we will describe the mathematical model developed to encode the applicable harmony rules, which will be responsible of generating the harmonizing chord from the operator's melody in the robotic experiments. The second part of the diagram in Figure 10 will be discussed in Chapter 5.

4.1 The Harmonizing Process

As a discipline of music theory, harmony deals with the simultaneous combination of pitches, namely *chords*, and their evolution towards other pitches over time. The process of harmonizing a melody line implies the election of the most suitable chords for accompanying such melody. During the musical piece, the chord heard at each moment has to be in consonance with the note that is sounding in the melody, but also has to be connected with the previous and following chords so that the overall composition has a sense of phrasing, of harmonic progression. In this section we introduce the core concepts necessary to understand how a melody line can be harmonized through musical chords.

4.1.1 The Key of the Harmonization

The first step for harmonizing a melody is to identify the key in which it is written, as the scale associated with that key will provide the notes with which the harmonizing chords will be constructed. The purpose of this section is to introduce the concept of *scale*, as its structure, given by the predefined distance between the notes, will serve as the generator of the harmonizing chords and the rules which are encoded at the end of this chapter.

A *scale* is simply a progression, ascending or descending, of musical notes from any note until its octave[2][50]. The importance of the scale is that it constitutes the basic structure used during the Common Practice Period: when creating a musical piece, composers often chose a scale as a frame for a musical piece, being the notes belonging to that scale the main construction elements for the creation of melodies and harmonies. Therefore, in order to harmonize a melody we must establish the key in which the melody is written, since the suitable harmonizing chords for the melody will be constructed using the notes of the scale associated with the key.

The structure of the scales is given by the distribution of the different intervals in the progression. Although an uncountable number of scale types can be encountered in the

[2] An *octave* is the interval between two musical pitches, where one of the pitches doubles the other in frequency.

traditional music all over the world, in this work we will focus on those used in the Western musical tradition. Scales in the Western music are conditioned by the division of the octave into 12 semitones, being a tone composed by two semitones. This division can be observed in Figure 11, where the 12 notes comprised within an octave are shown in a piano keyboard. Because of this discretization of the octave, the semitone constitutes the unit of measurement in the context of harmony and it will allow the definition of the different scale and chord types, in which our formation control strategy will be based.

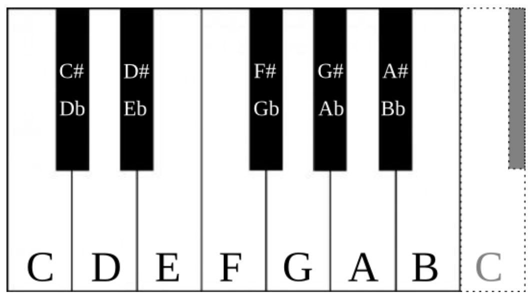


Figure 11: Interval division of the octave according to the European system. The distance between two consecutive white keys is one tone if there is a black key in between, and a semitone if there is not (steps E-F and B-C). The distance between a white key and an adjacent black key is one semitone.

In general, four types of scales are considered: pentatonic, whole tone, diatonic and chromatic [51]. In this work, we will focus solely on the diatonic scale as it is the foundation of the Western musical tradition and base for the harmony rules we study. The diatonic scale contains seven pitches per octave organized in five tone steps and two semitone steps [51]. The most popular modes of the diatonic scale are the ionian and the aeolian, which correspond to the major and natural minor scales, respectively. We will focus in this work on harmonizations of major keys, whose interval pattern between notes is depicted in Figure 12.

Now that we have established the structure of the major diatonic scales and its importance in the harmonization process, we need to determine the procedure to establish the key of the harmonization in our robotic context. In general, the key of a piece is determined by analyzing its initial and final chord, since both usually correspond to the tonic chord of the

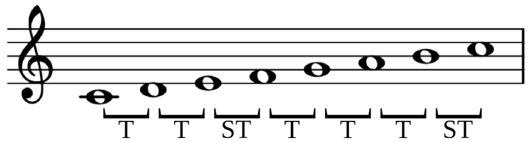


Figure 12: Interval pattern of a major diatonic scale. In particular, the depicted scale is the C major scale, which results from playing all the consecutive white keys in the piano extract depicted in Figure 11.

musical key (e.g. if a composition is written in the key of G major, there is a high probability that the first and last chord of the composition are the G major chord). However, when generating multi-robot formations of a particular geometry from the input melody, we do not have the complete melody in advance. Instead, the harmonies are generated as a function of the input at each moment. Therefore, we need to come up with a strategy to define the music key. In our application, we have chosen to identify the first note played by the operator as the tonic, and use the major scale associated with that tonic to create the harmonies.

4.1.2 Using the Suitable Scale Degree

We have dedicated the previous section to the scale, frame for a musical piece and basic structure used in composition during the Common Practice Period. The importance of scales relies on the structures that can be constructed with their notes, the *chords*, which we will apply to the operator's input notes to harmonize them.

Once the key of the composition has been identified, we can construct set of seven triads that we need for harmonizing the melody. We obtain those triads by stacking on top of each note the corresponding third and fifth note, in ascending order, as depicted in Figure 13. Due to the interval structure of the diatonic scale, these triads have a very particular structure in terms of distances between their notes, which will determine the type of resulting chord. In Chapter 5 we will describe the structure of these triads and how this structure will be mapped to the formation geometry to be displayed by the team.

At this point, we are in possession of the chords that can be applied to the melody

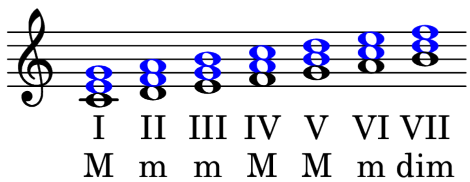


Figure 13: C major scale. The notes of the ascending scale are depicted in black and the triad formed over each of the notes is shown in blue.

input, harmonizing it. The simplest harmonizing approach is to use for each note of the melody a triad that contains it. For instance, when harmonizing an E note within the C major key, we could use the chords labeled I, III and V in Figure 13, as they all contain the note E. For the robots, every time the user presses a key, this would result in three potential formations to be adopted. But, can we use any of the three chords interchangeably for a given input? Certainly not, since they embody different musical functions within the composition. The structure of the diatonic scales generates the concept of tonality, which establishes the relationships between the notes in a scale and limits the set of acceptable chord combinations.

The tonal system is an organized system of tones in which one tone becomes the central point to which the remaining tones are related [52]. This central point is the *tonic*, which represents the complete relaxation, the target toward which other tones lead. The tonal system emerges as a result of the structure of the diatonic scale, and establishes the importance of certain notes over others. In a diatonic scale, the tonic is the note that gives the name to the key and has associated the biggest harmonic weight (for example, in the scale of A minor, the most important note is A). Because of their different harmonic roles, each note in a diatonic scale is identified with a *scale degree*. In Figure 14 the names of the different scale degrees are shown for C major. The use of roman numerals constitutes an alternative notation for the scale degrees.

The fact that the three potential chords suitable to harmonize an input note are not interchangeable arises from the different harmonic roles that each of the notes of the scale performs in the tonal scheme define above. This role relies on the tension associated with

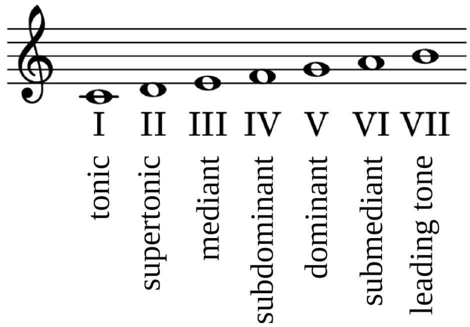


Figure 14: Scale degrees according to the tonal system.

the chord and its tendency to be resolved towards the tonic chord, which represents the maximum stability. Table 1 presents a common classification for the different scale degrees, as considered in [53, 54]. From this table the main functional behaviors in harmony can be deduced:

Tonic. Represents the tonal center of the scales, and its use is associated with the final resolution of the harmonic progression, which results in total relaxation.

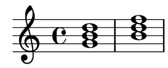
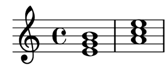
Dominant. Performs the opposite role to the tonic, as the chords over the V and VII degrees embody the maximum tension and need to resolution towards the tonic chord. The dominant function aims to create instability so that the ear claims the tonic for resolution.

Predominant. As it name suggests, these chords normally appear before a dominant chord and evolve towards it. They expand away from the tonic, creating instability that leads to the dominant and adding variety to the progression.

Table 1 also lists the harmonic function for the third and sixth degrees. Usually, these chords are used to prolong the sound of the tonic, adding variety to the progression, and preparing for the introduction of a predominant chord.

As a result of the tonal system associated with diatonic scales, the appropriateness of the different harmonic evolutions between chords have been encoded in harmony treatises (e.g. [53]). The thorough study of compositions from the Common Practice Period has

Table 1: Harmonic roles of the different scale degrees.

Harmonic function	**Degree**	**C major example**
Tonic	I	
Dominant	V, VII	
Predominant	II, IV	
Tonic substitution/extension	III, VI	

allowed the creation of Table 2, where the tendency of each scale degree of evolving into another has been depicted.

In conclusion, the tonal system determines the suitability of each of the three potential chords that could be used to harmonize an input note, due to the different roles of the notes of a diatonic scale. The chosen chord does not only have to contain the melody note to be accompanied, but also has to be connected with the previous and following chords so that the overall composition has a sense of tonal harmonic progression.

Table 2: Advisability of the different harmonic progressions between scale degrees [53]. The column labeled as *often* represents the most natural evolution of each scale degree, while the columns labeled as *sometimes* and *rarely* appear occasionally in a composition. Although the transitions in the column *rarely* may seem unnatural, they help introduce variations in a composition, avoiding excessive repetition.

Degree	**Often**	**Sometimes**	**Rarely**
I	IV, V	VI	II, III
II	V	VI	I, III, IV
III	VI	IV	II, V
IV	V	I, II	III, VI
V	I	IV, VI	II, III
VI	II, V	III, IV	I
VII	III		

4.1.3 Connecting the Voices

The final step on the harmonization process is to connect the notes in two consecutive chords. The inherent tension and relaxation found in the tonality structures described in Section 4.1.2 are responsible for the resulting phrasing of the harmonic progression. However, the harmonic progression does not completely characterize the evolution of the musical lines, since a triad is defined by the notes that it contains and not by the spacial arrangement of those notes in terms of pitches. Therefore, for the same harmonic progression two chords can be linked very differently. This section focuses on the harmony rules taken into account when connecting the notes of two chords.

Compositions are often based on the use of four voices that try to emulate the different registers of human voices: soprano, alto, tenor and bass [53]. As an imitation to human capabilities, the pitches that each of these voices can achieve is limited, as shown in Figure 15.

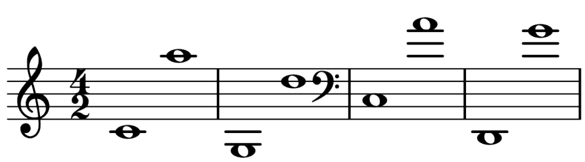
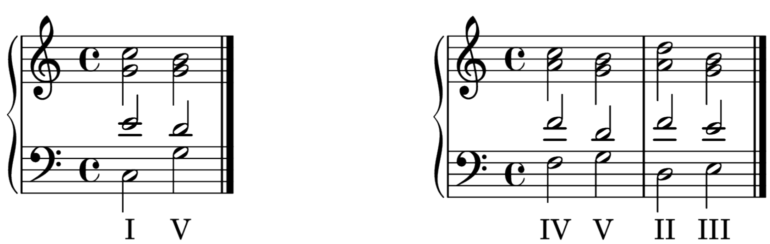


Figure 15: Allowed ranges for each of the voices in composition.

A common practice, adopted in this work, is to build the harmonies on top of the bass, as this voice contains a much bigger harmonic weight if compared with the others. Given its higher harmonic hierarchy, we identify the input melody provided by the user with the bass line, creating the harmonizing triads with the other three voices by placing one note of the triad in each of them. Thus, the bass corresponds with the goal locations commanded to the leader, whereas the notes that harmonize the bass represent the formation to be depicted by the followers, as they depend on the objectives of the leader.

There exist many rules regarding chord connection and voice leading, which aim for a smooth horizontal evolution in each of the voices. A subset of these indications gives advice to the composer about how to move each of the voices to achieve soft and pleasant harmonizing movements. However, a bigger set of these rules refers to forbidden practices that must be avoided under any circumstance. In this section we include a summary of some of the rules taken into account during this work, extracted from [53]. This summary does not intend to include all the possible combinations subject to be encountered when harmonizing a melody line, but to illustrate the main situations that had to be considered throughout the work presented in this thesis.

In order to obtain a smooth horizontal evolution in each of the musical lines, each note should move to the nearest available position, avoiding any unnecessary jumps. When connecting triads with a given bass in root position, if there exist common notes between the two chords, this objective translates into repeating those common notes in the same voice, while the remaining voices move to the nearest available position. This situation can be observed in Figure 16a, where the note G in the alto voice is common to the C major and G major chords. However, the triads to be connected need not have common notes. Figure 16b presents two cases where this situation can be observed, which is overcome by moving the three upper voices to the nearest available position, in the opposite direction to the bass movement.



(a) Connection of triads with common notes.

(b) Connection of triads with no common notes.

Figure 16: Examples of connection of triads in root position.

In addition to this indications, we also took into account several prohibitions regarding movements that need to be avoided. In counterpoint instruction, the use of intervals like fifths and eights often results in undesirable situations due to the auditive repetition and dullness that these movements produce. The parallel motion between chords is strictly forbidden between chords, and the direct motion (when two or more voices move in the same direction and, although the first interval between them is not a fifth or eigth, the resulting interval is one of the two) is generally forbidden, with exceptions [53].

In conclusion, the rules presented in this section aim for a smooth horizontal movement when connecting triads in root position. This summary aims to establish a general basis of the rules taken into account when creating the mathematical model for the harmonization process. For an exhaustive collection of applicable rules, please see [53].

4.2 Encoding Harmony Rules

The harmony generation process depicted in Figure 10 as the first phase of the human-swarm interaction has been described throughout this chapter. The harmonization process detailed in Section 4.1 in synthesized in the mathematical model presented in this section. The model developed in this work is based in the use of two automata: one is in charge of determining the scale degree to transition to based on the tonal system presented in Section 4.1.2, while the other determines how the voices between two consecutive chords are connected, according to the voice leading rules of Section 4.1.3.

We have restricted the harmonizing rules to those which apply for the connection of triads in root position, which is a common simplification in beginner practices of harmony. Using the root position of a triad implies that the fundamental note or *root* (depicted in black in Figure 13) must be located in the bass line. Therefore, the note the operator plays in the keyboard unequivocally determines the scale degree to be harmonized, simplifying the implications of the tonal system. The scale degree transitions resulting of this approach have been encoded in an automaton depicted in Figure 17. The states of this automaton correspond to the scale degree that will be used in the harmonization, and its transitions are triggered by the human user's input. The actual values of this automaton's transitions are instantiated with the first input of the operator, note that establishes the musical key of the

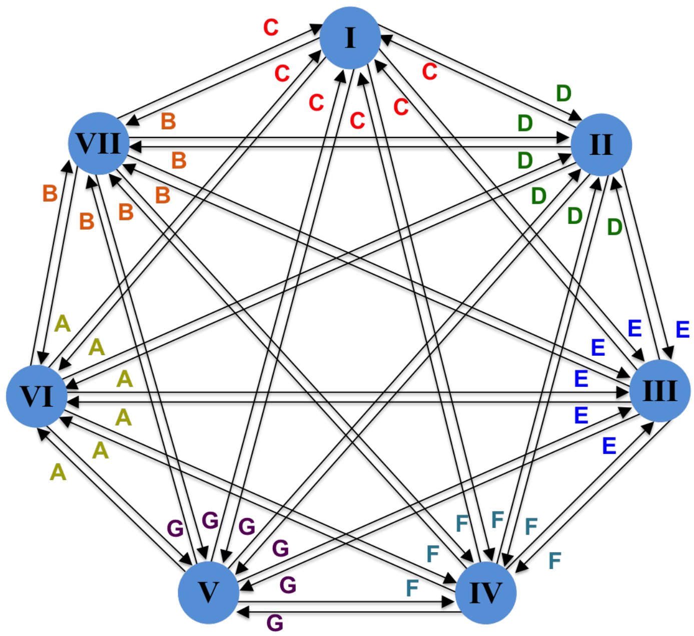


Figure 17: Leader's automaton for the C Major scale under the simplification of harmonizing using only triads in root position. The states of this automaton correspond to the scale degree to be harmonized by the followers, while the transitions between scale degrees correspond to the notes played by the operator.

harmonization, as explained in Section 4.1.1.

The harmony rules for smooth voice leading described in Section 4.1.3 have been encoded in the automaton shown in Figure 18, which for clarity purposes has been instantiated for the C major scale. The states correspond to the set of notes in the scale, available to the voices. When the first note of the melody is played in the keyboard, establishing the key and the scale, the states of the automaton are instantiated with the notes in the corresponding scale and the three states associated with the tonic chord are initialized. The transitions between states are triggered by the leader's automaton, whose evolution between scale degrees depends on the new notes introduced by the human operator through the piano keyboard. The output of the automaton in Figure 18 is a series of chords, one per each note in the input, whose structure will be used for the selection of the formation geometry, as we describe in the upcoming section.

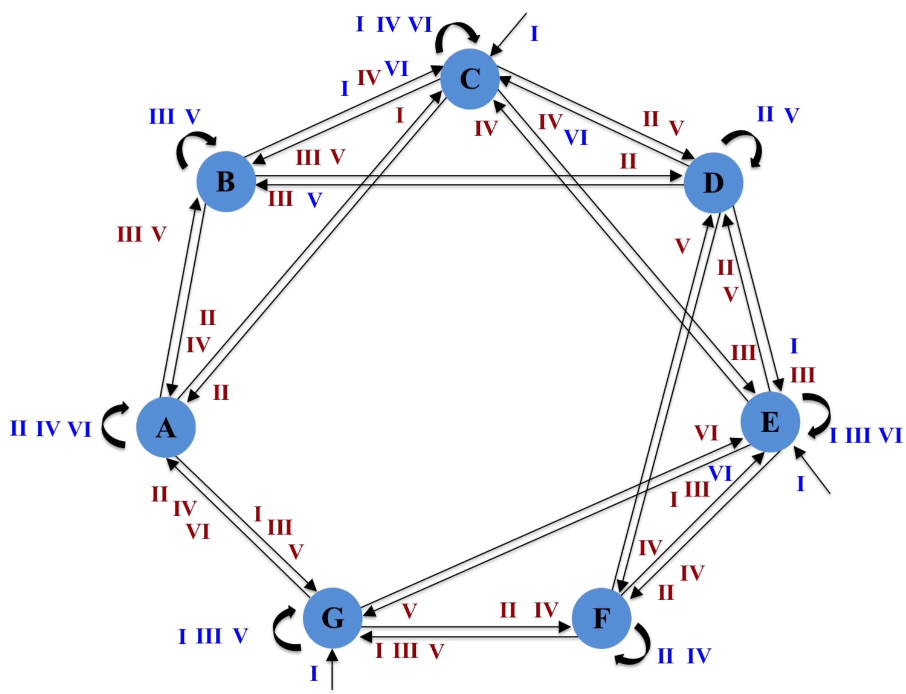


Figure 18: Follower's automaton for the C Major scale created using voice leading rules for triads in root position[53]. Transitions depicted in maroon indicate a dependency with the previous transition, while blue transitions are independent of previous commands.

CHAPTER 5

MAPPING TO ROBOT MOVEMENT

The novelty of the human-swarm interaction presented in this work is the use of music theory as a generator of formation geometries from a simple user's input. Figure 10 depicts the high level operation of the interaction strategy. In Chapter 4 we dealt with the first task in the flow chart: the automatic generation of the applicable harmonies to the user's input. This chapter focuses on the formation description part, that is, how do we use the chords obtained through the harmonization process to coordinate the group of robots. First, we will describe the criteria applied to the harmonizing chords to select the formation geometries. Next, we will define the control laws to be executed by the robot team in order to travel to the objective locations specified by the user while displaying the right formation.

5.1 Geometry Selection

The application of music theory to the user's input has allowed us to obtain complex structures from a simple input specification. For each user input, the output of the harmony generator designed in Section 4.2 is a chord, whose structure which will be used as a basis to generate the formation geometry to be depicted by the followers around the leader.

So far, all the musical elements and harmony rules we have introduced in Section 4.1 are associated with diatonic scales, characterized by a very particular pattern distances between the notes in the ascending progression, with five tone steps and two semitone steps. As a consequence of this uneven interval distribution between the notes, illustrated in Figure 12, the triads formed over the diatonic scale in Figure 13 have different structures according to the distance between their notes.

Triads are a particular type of chords, constructed by the superposition of two thirds. There exist four types of triads, according to the interval distances between the three notes of the chord: major, minor, diminished and augmented. In particular, in major scales only

Table 3: Types of triads constructed over a major diatonic scale. The interval structure represents the distance between the fundamental and the third, and between the third and the fifth. *M3* stands for major third (distance of 2 tones) and *m3*, for minor third (1 tone and 1 semitone).

type of triad	major	minor	diminished
interval structure	M3 & m3	m3 & M3	m3 & m3

the first three types are encountered (the augmented chord appears only in certain modes of minor scales). The interval structure of each of this type of chords is depicted in Table 3. For our multi-robot control strategy we will analyze the output chord produced by the automaton in described in Section 4.2, and map its type, that is, its interval structure, into a predefined formation geometry. This way, when the human operator plays a new note in the keyboard that directs the leader robot, the followers will depict a formation whose structure will illustrate the type of the chord associated with the human input within the selected harmonizing key.

5.2 Mapping to Robot Movement

The election of a leader-follower paradigm for the movement of the robots was tightly related to the interaction strategy depicted in Figure 10. This configuration gives us ability to direct the group to locations of interest by setting goals for the leader at the same time that we can use the structured outcomes produced by the musical generator as a selection criterion for the formation's geometries. In addition, the inner nature of the composition process, where the evolution of the harmonizing voices is totally subordinate to changes in the melody, also relates to a leader-follower strategy.

The purpose of this section is to translate the geometry specification obtained from analyzing the chord type provided by the harmony generator and transform it into feasible commands executable by a multi-agent system. We will describe how to the single integrator dynamics are obtained under the specification of a formation geometry. We also include a section the assignment algorithm implemented in the robotic experiments presented in the next chapter with the objective of minimizing the total distance traveled by the agents.

5.2.1 Control Laws for the Robot Team

The type of the chord produced by the harmony generator has chosen to be the selection parameter between different formations to be displayed by the robot team. For a major scale, considering the only use of triads, a total of three possible formation geometry outcomes will be potentially displayed by the team, according to Table 3.

Let's consider a total of N planar robots in our team, with one designated leader and $N-1$ followers. As introduced in Chapter 3, for each of the possible formations to be displayed by the team we can define the geometry of the formation in terms of distances between the agents in the team. Therefore, for each formation we will have a set D of relative, desired interagent distances:

$$D = d_{ij} \in \mathbb{R} \quad | \quad d_{ij} > 0, \quad i, j = 1, ..., N, \quad i \neq j \tag{23}$$

For simplicity, we will assume that the velocity of each robot can be controlled using single-integrator dynamics, which we will transform into the unicycle robot dynamics in the implementation presented in Chapter 6 using a near-identity diffeomorphism [55]. The assumption of single-integrator dynamics allows us to write the single-integrator dynamics as:

$$\dot{x}_i = u_i, \quad i = 1, \cdots, N \tag{24}$$

with $x_i \in \mathbb{R}^2$ the position of the robot and $\dot{x}_i$, its velocity.

We begin defining the control law for the leader's robot. The objective of this robot is to travel to the target position in the workspace corresponding to the current user's input. Thus, the dynamics of this robot are simply defined by:

$$\dot{x}_l = x_g - x_l \tag{25}$$

with x_g the location of the goal, communicated to the leader robot by the control system, and l the index of the robot selected as a leader.

In order to define the control law of the followers, we must take into account that the usual situation in a leader-follower strategy is the lack of global information by the follower agents. Generally, robot i does not possess information about absolute positions in the workspace. Data like its absolute position, x_i; its neighbors' absolute position, x_j, $j = 1, \cdots, N$, $j \neq i$; or the goal location, x_g; are unknown for a follower robot. Instead, the information available to this robots is local, in terms of relative distance of those external elements. With these assumptions, achieving a formation specified by D is feasible defining the follower's controller through a variation of the consensus equation [49]:

$$\dot{x}_i = \sum_{\substack{j=1,\ldots N \\ j \neq i}} w_{ij}(x_j - x_i) \tag{26}$$

with $w_{ij} = \frac{\|x_i - x_j\| - d_{ij}}{\|x_i - x_j\|}$ (see Chapter 3). Note that although the terms x_i, x_j that appear in Eq. 26, they do not imply the use global information by the follower robot. Since they are presented in differences, these terms indicate relative positions of the neighbors with respect to the agents local coordinate system.

5.2.2 Assignment Algorithm

When we decided to map each type of chord into a certain predefined formation, we left out the consideration of any nuances associated with the note played by each voice. This means that the objective for the followers is to create a shape that depends on the resulting triad, but the role of each follower is not specifically connected with any of the notes of the chord. Since we made this simplification, we decided to assign the roles of the followers to the formation positions so that the total distance traveled by the agents to achieve the formation is minimum.

The assignment problem to be solved is denominated the problem of parametrized assignment, with certain variations. In the original problem a team of mobile robots must decide what role to take on a given planar formation, taking into account the rotation and translation of the formation [56]. However, in our application the rotation of the formation is allowed, but not the translation. The latter is conditioned to the position of the leader. In our problem, $y_1, y_2, ..., y_{N-1} \in \mathbb{R}^2$ represent the planar positions of the targets in the formation. Let's define $x_c \coloneqq \frac{1}{N}\sum_{i=1}^{N-1}$ as the centroid of the followers position and y_c, will be the centroid of the target positions, which in this case corresponds to the position of the leader robot. We can rewrite the positions of agents and targets with respect to these centroids as:

$$x_{i_c} = x_i - x_c \tag{27}$$

$$y_{i_c} = y_i - y_c \tag{28}$$

$\forall i = 1, \ldots N-1$. We also denote the rotation of the formation as $\theta \in [0, 2\pi)$ and the assignment of the agents to targets as p, which is an element of P_{N-1}, the set of all possible permutations over $N-1$ elements. The notation $p(i)$ will represent the i-th element of p. We can now define our optimization problem:

$$\Sigma_{l_2^2}(x, y) : min_{[0,2\pi)\times P_{N-1}} \sum_{i=1}^{N-1} ||x_{ic} - R(\theta) y_{p(i)c}||^2 \tag{29}$$

with $R(\theta)$ as the rotation matrix

$$R(\theta) \triangleq \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{30}$$

The problem of solving the two parameters simultaneously – the angle, θ, and the permutation, p – is nontrivial. We used one of the suboptimal heuristic solutions in [56], named as *improved angular discretization*.

The first step of this heuristic approach is to find the optimal solution $(\bar{p}, \bar{\theta})$ to the problem

$$\min_{\theta \in \hat{\theta}_0, \hat{\theta}_1, \ldots \hat{\theta}_{d-1}} \min_{p \in P_{N-1}} \sum_{i=1}^{N-1} ||x_{ic} - R(\theta) y_{p(i)c}||^2 \tag{31}$$

Once this first result is obtained, the authors propose a further step that enhances the results and brings it closer to the optimal solution:

$$\begin{cases} p^{\#} \triangleq p^{*}(\bar{\theta}) \\ \theta^{\#} \triangleq \theta^{*}(p\#) \end{cases} \tag{32}$$

being $(p^{\#}, \theta^{\#})$ the suboptimal solution used for the assignation of targets.

At this point, we have proposed a human-interaction approach based on music theory abstractions that is able to interpret a note of a melody within a tonal framework, identify the most suitable harmonizing chord within the corresponding key, use the interval structure of this chord to select the representative formation control geometry and execute it around a robot whose aim is directly commanded by the melody. In the next chapter we will describe the experiments carried out for this thesis, which blend all the elements introduced in Chapters 4 and 5.

CHAPTER 6

EXPERIMENTAL WORK

The human-swarm interaction modality presented in Chapters 4 and 5 has been implemented in two different multi-robot systems in order to illustrate the effectiveness of the proposed method across platforms. In each of the experiments, a team of four differential drive robots were deployed in the platform workspace, where the target locations were identified by points labeled with the notes' names. The human user was given the ability of injecting controls to the robotic team by specifying a melody on an electronic piano keyboard.

This chapter begins with the description of the experimental setup of both experiments included in this work, where we include the details about the two robotic platforms, the selected formation geometries and the visualization of the workspace. In the latter part of the chapter, we describe the outcome of the robotic experiments.

6.1 Experimental Setup

The execution of the proposed human-swarm interaction algorithm on a robotic platform implies certain requirements to be met by the robotic platform. First, there must be available a team of mobile robots capable of moving independently in the platform workspace. Second, the platform needs a sensor system capable of providing the robot team information about the target location to be attained, together with the relative positions of other robots in the team. Lastly, the platform needs to have a device that allows the human operator to interact with the robot team, in order to introduce the operator's commands.

In this section we describe the main characteristics of the two robotic platforms in which we carried out the experimental work presented in this thesis. We have also included the transformation from single integrator dynamics to unicycle dynamics to allow the implementation of the control laws defined in Section 5.2.1, together with the definition of the formation geometries to be depicted by the robot team and the location of the goal positions to be visited.

6.1.1 Robotic Platforms

Two different platforms were used to illustrate the effectiveness of the human-swarm interaction modality introduced in this thesis. The first robotic platform comprises a team of four Khepera III robots from K-team, which were deployed in the platform's arena, a planar surface surrounded by a motion capture system, Optitrack, capable of providing information about localization. The other platform used in the experimental section of this thesis is the Robotarium, a remote-access testbed, where the algorithm is run on a team of GRITS-Bots. In both cases, the device provided to the human user was a electric piano keyboard, whose input can be fed to the computer program in charge of running the algorithm.

6.1.1.1 Khepera III

The Khepera III robot is a small robotic platform specifically designed for robotic swarm-type experiments [57]. The design of the Khepera III, depicted in Figure 19, is that of a differential drive robot, with two independently driven wheels located at both sides of the robot and a sliding caster for stability. The sensor equipment of the robot includes embedded encoders in the two wheels, which allow a precise control of their velocities; together with infrared and ultrasonic sensors for short and long range object detection, respectively. The Khepera III robots each have a 600 MHz ARM processor with 128Mb RAM, embedded Linux and a wireless card for communication over a wireless router, which allows for receiving commands and processing commands from a remote machine.

The ultrasonic and infrared detectors were not used in this experiment. Instead, the data relative to the position of the robots was provided by ten OptiTrack S250e motion capture cameras [58]. The information about the position and orientation of each agent was obtained by the tracking system, which distinguishes between different robots thanks to the distinct pattern of the reflective markers located on top of the robots, as shown in

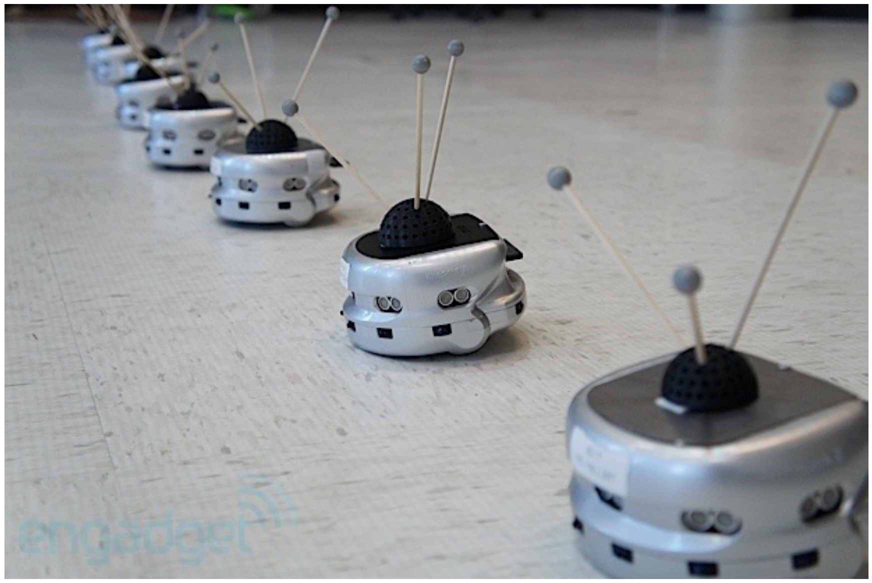


Figure 19: Team of Khepera III robots equipped with reflective markers subject to be detected by the Optitrack tracking system.

Figure 19. Although the tracking system offers information about the global position of the robots, the controller of each robot was fed the relative positions of its neighbors by the control program. In the case of the leader robot, the controller was also fed the relative position of the target to be attained.

The human user was provided a Casio CTKVK3 piano keyboard to control the swarm, as was shown in Figure 1. The keyboard's MIDI[1] output was continuously sent to the control system throughout the experiment, allowing real time control the group of robots with the piano keyboard. Every time the operator pressed a key, a new note was introduced in the melody line and the harmonizing chord was produced according to the automata in Figures 17 and 18.

6.1.1.2 The Robotarium

The Robotarium is a remote-access testbed that provides researchers from all over the world with the capability to develop and test multi-agent applications [59]. The Robotarium is

[1]Musical Instrument Digital Interface.

populated by the GRITSBots [60], a smaller and lower-cost robotic platform when compared to the Khepera III setup described in the previous section.

Figure 20: The Robotarium, a remote-access testbed.

The GRITSBot, depicted in Figure 21, has a kinematic design of a differential drive robot, with a very small footprint: its dimensions are approximately those of a 3cm cube, 77% smaller than the Khepera III. The GRITSBot is equipped with a main processor that handles the received commands via WiFi. An independent processor manages the wheels' stepper motors, so that the use of encoders is not needed in the GRITSBot design.

Analogously to the experiment setup described in Section 6.1.1.1, the interaction device in this platform was again an electronic piano keyboard, as shown in Figure 25a.

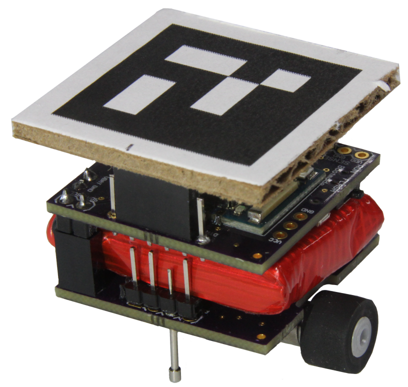


Figure 21: A GRITSBot with the ArUco tag on top for tracking.

6.1.2 From Single Integrator to Unicycle Dynamics

Both robotic platforms described in Section 6.1.1 consist of differential drive robots, whose locomotion is produced by two independently driven wheels located at opposite sides of the robot's body. Due to its kinematic design, the robot can travel between any two positions in the plane, but it cannot follow any trajectory since it is not a holonomic configuration (e.g. it cannot move in the direction of the axis of the wheels).

In Section 5.2 we treated the robot dynamics as those of a single integrator: a point capable of moving in any direction in the plane. Therefore, we must transform the obtained control law defined in terms of a single integrator to control commands executable by a differential drive robot. The desired velocities in terms of a single integrator dynamics were given in Eq. 26 and 25, where x_i and $\dot{x}_i \in \mathbb{R}^2$ represent the position and velocity of robot i. We can obtain the unicycle dynamics using the $\dot{x}_i$ term calculated in the control law as:

$$v_i = [\cos(\theta_i), \sin(\theta_i)] \cdot \dot{x}_i, \quad \omega_i = \arctan\left(\frac{[-\sin(\theta_i), \cos(\theta_i)] \cdot \dot{x}_i}{[\cos(\theta_i), \sin(\theta_i)] \cdot \dot{x}_i}\right) \tag{33}$$

where θ_i is the heading of the robot and v_i and ω_i are the linear and angular velocities, the inputs for the unicycle robot. These velocities are finally mapped to wheel velocities, ω_L

and ω_R, through the expression:

$$\begin{bmatrix} \omega_L \\ \omega_R \end{bmatrix} = \begin{bmatrix} \frac{2v-\omega d_B}{d_R} \\ \frac{2v+\omega d_B}{d_R} \end{bmatrix} \tag{34}$$

where d_R represents the wheel diameter and d_B the distance between the two wheels, the robot base.

6.1.3 Workspace Visualization

The human-swarm interaction approach presented in this thesis relies on the leader-robot paradigm, where the human operator is given the ability of specifying the target locations to be visited by the leader robot. Throughout this document, we have not specified what those targets ought to be since they depend of the particular application to which this algorithm is applied.

In the experimental work of this thesis, the instantiation of the target locations for the robot team was necessary. The goal positions for the leader of the robot team were not designated at random within the workspace. Instead, we chose a musically-inspired geometry to locate the targets: the chromatic circle. The chromatic circle represents the relationships among the twelve equally-space notes present within an octave, illustrating the cyclic distances existent between pitches [51]. The positions to be visited by the leader robot were located at the extremes of twelve equidistant radii of the chromatic circle, as represented in Figure 22. In both experiments, an overhead projector was used to project this spacial representation over the physical environment to provide a visualization to the objective positions along with their associated note to the operator.

6.1.4 Formation Control

In each of the experiments, we used a team of four robots, aiming to represent the four voices commonly used in composition: soprano, alto, tenor and bass. For each note introduced by the human user, the type of the resulting harmonizing chord was chosen as the criterion to switch between different formations. As we mentioned in Section 5, the harmonic instantiations considered in this work were based on major scales, which results

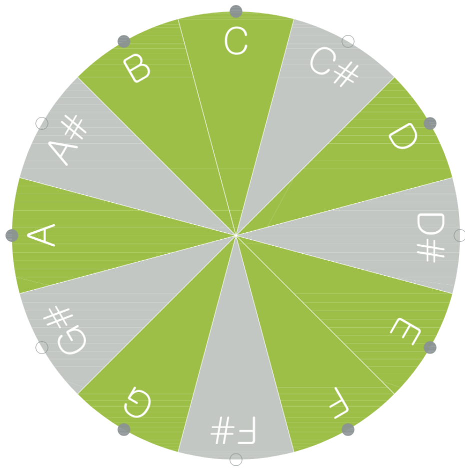


Figure 22: Chromatic circle used to represent the goal positions for the leader robot. The sectors corresponding to the C Major scale were highlighted as it was the key used for the experiment.

in a limited number of chords types subject to be encountered, namely major, minor and diminished (see Table 3).

The formations chosen to illustrate each of these chords, depicted in Figure 23, were triangles. In each of the formations, the centroid of the triangle (in red in the figure) is occupied by the leader robot due to its importance in the formation of the chord. The three vertices of the triangle represent the three notes that comprise the triad that harmonizes the note played by the human user. The different colors in the formations were used for visual feedback purposes.

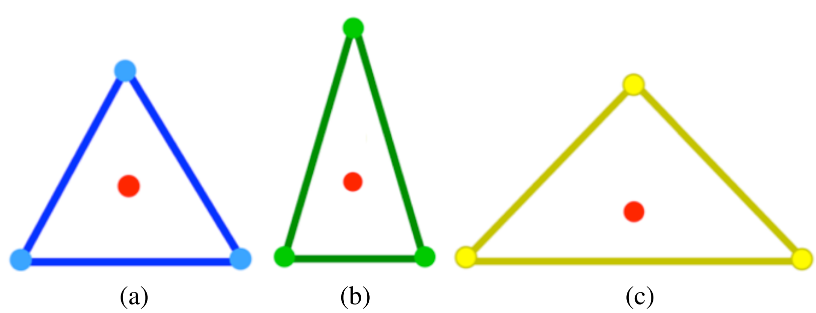


Figure 23: Selected formation geometries for the followers. (a) depicts the formation for major chords, (b) for diminished chords, and (c) for minor.

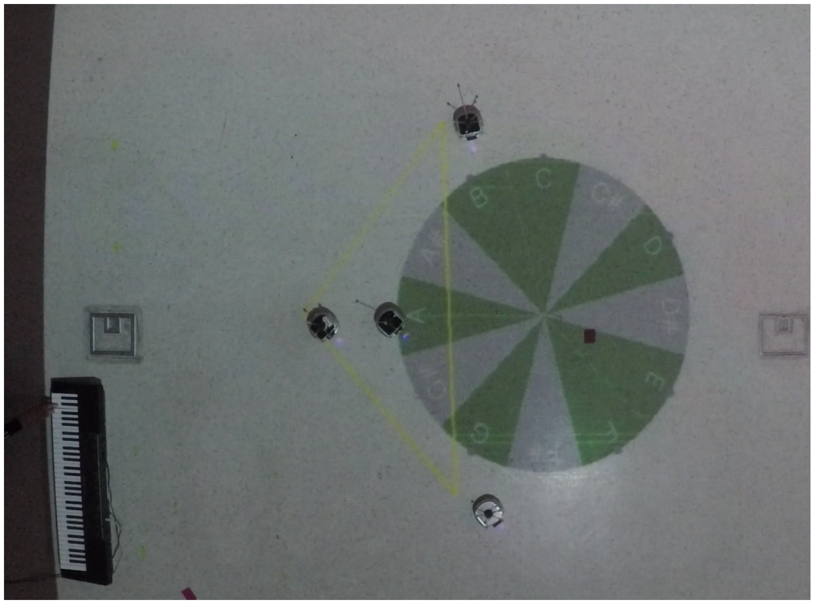
6.2 Experimental Results

The performance of the human-swarm interaction approach implemented in this thesis is detailed in this section for both robotic platforms described in Sections 6.1.1.1 and 6.1.1.2.

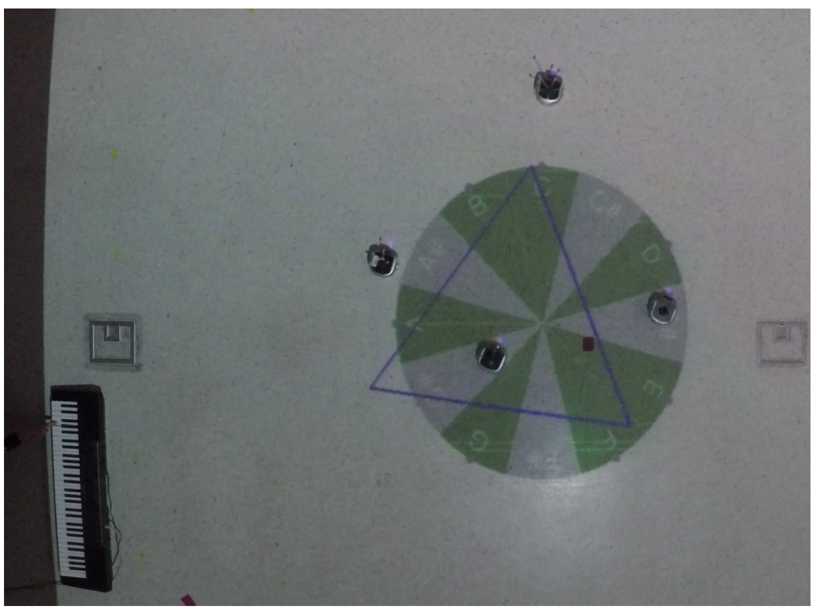
The behavior of the implemented control strategy for the Khepera III platform is depicted in Figure 24. In both pictures, the human operator is located in the bottom left corner of the picture with the piano keyboard. Figure 24a represents a situation where the robot team has successfully responded to the human input: the leader robot, located in the center of the projected triangle, has arrived to the target note position (A), and the followers are situated in the vertices of the triangle, depicting the minor formation around the leader. Figure 24b illustrates an intermediate situation, where the leader is still traveling towards the target position, corresponding to G. The followers go after the leader, trying to achieve the desired major formation, depicted around the leader.

On the other hand, Figure 25 illustrates the operation of the experiment in the Robotarium. As described in Section 6.1.1.2, in this experiment the user is given again an electric piano keyboard as the device interaction, through which the operator can specify the location to be attained by the team, while supervising their global performance, as depicted in Figure 25a. Figure 25b presents a situation where the robot team have successfully arrived to the objective location specified by the operator and the leaders perform the correct formation around the leader robot.

The main conclusion observed in these experiments relates to the absence of variation for the same input command. For each of the notes to be played by the human operator, the resulting chord type remains invariant. This is a result of the simplifications adopted with respect to the available elements for harmonizations: the exclusive use of triads in root position restricts to one the set of possible scale degrees that can harmonize a note in the bass.

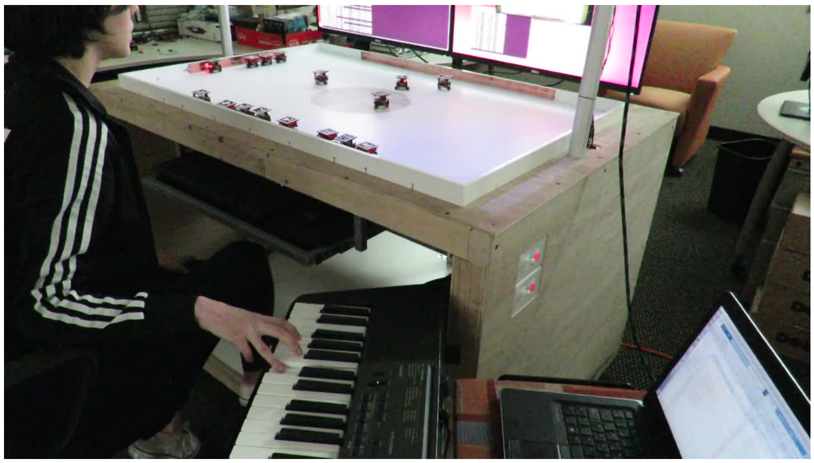


(a) Followers depicting a minor chord formation around the leader's goal.

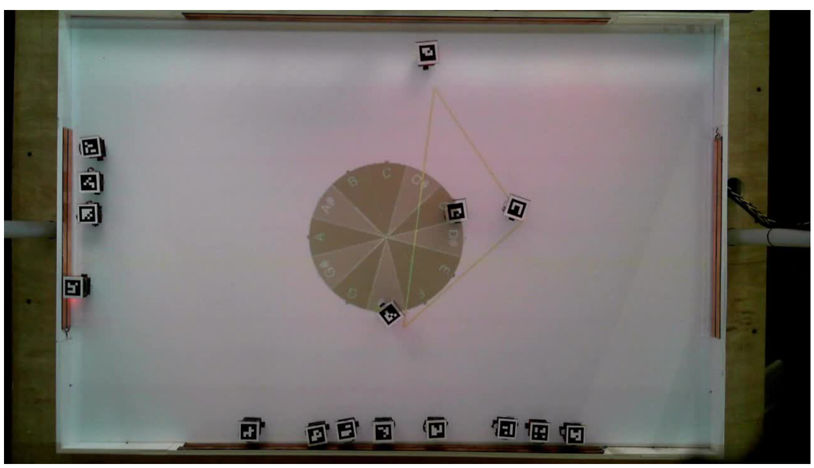


(b) Leader robot heading to its goal while the followers try to maintain the corresponding formation while following the leader.

Figure 24: Multi-robot implementation of proposed human-swarm interaction strategy on the Khepera III platform.



(a) User introducing input commands through the piano keyboard.



(b) GRITSBots depicting the minor chord formation corresponding to the D minor chord.

Figure 25: Implementation of the proposed human-swarm interaction on the Robotarium.

CHAPTER 7
CONCLUSIONS

The increasing number of multi-robot applications has raised the importance of the topic of human-swarm interactions. However, while potential benefits are subject to arise when incorporating a human operator in the control of a multi-robot system, consensus has not been achieved about about how the high level information provided by the human user can be injected into the multi-robot control strategy.

We have explored in this thesis a new human-swarm interaction modality where music theory plays the leading role, serving as the abstraction to model the input commands of the human operator into the team's control laws. We have focused on harmony, an extensively studied field of music theory, as a way to modify the control parameters of the swarm. A leader-follower strategy has been chosen to illustrate the music-based control, where the piano input specifies the locations to be visited by the leader robot and the harmonies determine the formation to be displayed by the team.

The main concepts involved in the process of harmonizing a single melody line have been presented. We have applied those concepts to produce a mathematical model that encodes harmony rules through automata. This model allows the system to automatically decide which scale degree it should transition to and generate the appropriate chord for that degree, complying with the applicable harmony rules. The output of these automata has been interpreted to produce robotic movement, using the type of the generated chord by the automata as the criteria to switch between different formation geometries. The formation control law designed for the robots has been optimized so that the total distance traveled by the agents to achieve the formation is minimum. The designed multi-robot control strategy has implemented in two real robotic platforms, where a piano keyboard has been used to direct the robot team towards areas of interest. The piano commands have been responsible of the automatic generation of the formation geometries as well. Although we have adopted a series of simplifications to mitigate the complexity of music creation, a satisfactory outcome has been obtained. In the next chapter we introduce further musical elements that could be incorporated into the presented interaction approach in future work.

In conclusion, the interaction approach introduced in this thesis has demonstrated the possibility of using music theory as a way of controlling multi-agent systems. The effectiveness of this interaction strategy has been tested in two robotic platforms. Through an intuitive input specification, the user has been able to dictate where the robot team should go and what its spacial arrangement should look like, results that fully support the viability of the proposed interaction strategy.

CHAPTER 8

FUTURE WORK

This thesis has introduced a novel human-swarm interaction approach based on harmony which allows the generation of formation geometries from a human operator's input. Music theory offers a great variety of tools that allow for the creation of very complex effects. However, since the purpose of this work was to explore the potential of using harmony rules as an abstraction for multi-robot control, we chose to limit the musical elements to be modelled within the extensive toolbox offered by harmony techniques.

Throughout this work, we have only considered harmonizations of the bass line with chords in root position. The most immediate extension of this work would be to remove this restriction, allowing the use of the three inversions of a triad, which would result in the possibility of choosing between three harmonizing chords when introducing a new control input through the keyboard. This approach would overcome the lack of variety observed in the experimental work, as we would abandon the one-to-one mapping in favor of transitions between scale degrees according to their importance in the tonal system. Therefore, considering different inversions of the chords will notably broaden the spectrum of harmonizing possibilities.

When harmonizing a melody, we need to know in which key that melody is written in order to select the appropriate chords to harmonize it. However, modulation is a very common practice in composition that consists in the change of the key a musical piece into another. We are interested in exploring how modulation could be applied to the injection of notes that do not belong to the selected key, since the change of key would affect how the notes are harmonized due to the role change produced by the tonal system.

The final aspect we would like to consider in the musical context is to change the role of the user's commands, considering the input melody as the soprano line instead of the bass. When we listen to a musical composition, we often identify the high-pitched notes as the melody of the composition and the lower voices as the accompaniment. However, considering this type of harmonizations is often more complicated than harmonizing the bass, since the latter carries much more tonal information than any other voice in the composition and the high-pitched melodies are often less steady than the deep ones. Therefore, introducing this change will probably result in a redefinition of the mathematical model and the consideration of group of notes to be harmonized instead of producing one harmonizing chord for each user's input.

The mapping aspect between the musical abstractions and the robot movement is also a subject to be extended. In this work, the chosen criterion to switch between different formations was the type of the resulting chord. We are interested in including other criteria in the formation control mapping, such as the specific configuration of the notes that constitute the chord or the harmonic role of the chord within the tonality. However, our main interest regarding the mapping from the musical abstractions to the formation control relies on the ability of automatically generating the formation geometry from the analysis of the generated harmony, rather than selecting between a catalog of predefined geometries. For that purpose, we need to study which musical characteristics are subject to being mapped into an euclidean geometry.

REFERENCES

[1] N. Agmon, S. Kraus, and G. Kaminka, "Multi-robot perimeter patrol in adversarial settings," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pp. 2339–2345, May 2008.

[2] L. Iocchi, L. Marchetti, and D. Nardi, "Multi-robot patrolling with coordinated behaviours in realistic environments," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pp. 2796–2801, Sept 2011.

[3] C. Amato, G. Konidaris, G. Cruz, C. A. Maynor, J. P. How, and L. P. Kaelbling, "Planning for decentralized control of multiple robots under uncertainty," *CoRR*, vol. abs/1402.2871, 2014.

[4] S. Hoshino, H. Seki, Y. Naka, and J. Ota, "Multirobot coordination for flexible batch manufacturing systems experiencing bottlenecks," *Automation Science and Engineering, IEEE Transactions on*, vol. 7, pp. 887–901, Oct 2010.

[5] J. Lin, A. Morse, and B. Anderson, "The multi-agent rendezvous problem," in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 2, pp. 1508–1513 Vol.2, Dec 2003.

[6] J. Cortes, S. Martinez, and F. Bullo, "Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions," *Automatic Control, IEEE Transactions on*, vol. 51, pp. 1289–1298, Aug 2006.

[7] W. Burgard, M. Moors, C. Stachniss, and F. Schneider, "Coordinated multi-robot exploration," *Robotics, IEEE Transactions on*, vol. 21, pp. 376–386, June 2005.

[8] S. Lee, Y. Diaz-Mercado, and M. Egerstedt, "Multirobot control using time-varying density functions," *Robotics, IEEE Transactions on*, vol. 31, pp. 489–493, April 2015.

[9] T. Balch and R. Arkin, "Behavior-based formation control for multirobot teams," *Robotics and Automation, IEEE Transactions on*, vol. 14, pp. 926–939, Dec 1998.

[10] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *Robotics and Automation, IEEE Transactions on*, vol. 17, pp. 947–951, Dec 2001.

[11] M. Ji and M. Egerstedt, "Distributed coordination control of multiagent systems while preserving connectedness," *Robotics, IEEE Transactions on*, vol. 23, pp. 693–703, Aug 2007.

[12] J.-P. de la Croix and M. Egerstedt, "A control lyapunov function approach to human-swarm interactions," in *American Control Conference (ACC), 2015*, pp. 4368–4373, July 2015.

[13] H. Kawashima and M. Egerstedt, "Approximate manipulability of leader-follower networks," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pp. 6618–6623, Dec 2011.

[14] J. Chung, "Cooperative control of uavs using a single master subsystem for multi-task multi-target operations," in *Robot Intelligence Technology and Applications 3* (J.-H. Kim, W. Yang, J. Jo, P. Sincak, and H. Myung, eds.), vol. 345 of *Advances in Intelligent Systems and Computing*, pp. 193–212, Springer International Publishing, 2015.

[15] N. Leonard and E. Fiorelli, "Virtual leaders, artificial potentials and coordinated control of groups," in *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, vol. 3, pp. 2968–2973 vol.3, 2001.

[16] M. Egerstedt, J.-P. dela Croix, H. Kawashima, and P. Kingston, "Interacting with networks of mobile agents," in *Large-Scale Networks in Engineering and Life Sciences* (P. Benner, R. Findeisen, D. Flockerzi, U. Reichl, and K. Sundmacher, eds.), Modeling and Simulation in Science, Engineering and Technology, pp. 199–224, Springer International Publishing, 2014.

[17] M. Goodrich, B. Pendleton, P. Sujit, and J. Pinto, "Toward human interaction with bio-inspired robot teams," in *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*, pp. 2859–2864, Oct 2011.

[18] P. Ranganathan, R. Morton, A. Richardson, J. Strom, R. Goeddel, M. Bulic, and E. Olson, "Coordinating a team of robots for urban reconnaisance," in *Proceedings of the Land Warfare Conference (LWC)*, November 2010.

[19] Y. Diaz-Mercado, S. Lee, and M. Egerstedt, "Distributed dynamic density coverage for human-swarm interactions," in *American Control Conference (ACC), 2015*, pp. 353–358, July 2015.

[20] T. Setter, H. Kawashima, and M. Egerstedt, "Team-level properties for haptic human-swarm interactions," in *American Control Conference (ACC), 2015*, pp. 453–458, July 2015.

[21] S. Nunnally, P. Walker, M. Lewis, N. Chakraborty, and K. Sycara, "Using haptic feedback in human swarm interaction," in *Handbook of Research on Design, Control, and Modeling of Swarm Robotics* (Y. Tan, ed.), Hershey, PA: IGI Global, 2015.

[22] M. Diana, J.-P. de la Croix, and M. Egerstedt, "Deformable-medium affordances for interacting with multi-robot systems," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pp. 5252–5257, Nov 2013.

[23] S. Chopra and M. Egerstedt, "Heterogeneous multi-robot routing," in *American Control Conference (ACC), 2014*, pp. 5390–5395, June 2014.

[24] P. Velagapudi, P. Scerri, K. Sycara, H. Wang, M. Lewis, and J. Wang, "Scaling effects in multi-robot control," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 2121–2126, Sept 2008.

[25] S. Bashyal and G. Venayagamoorthy, "Human swarm interaction for radiation source search and localization," in *Swarm Intelligence Symposium, 2008. SIS 2008. IEEE*, pp. 1–8, Sept 2008.

[26] A. Kolling, K. Sycara, S. Nunnally, and M. Lewis, "Human swarm interaction: An experimental study of two types of interaction with foraging swarms," *Journal of Human-Robot Interaction*, vol. 2, pp. 103–128, June 2013.

[27] G. Podevijn, R. OǴrady, and M. Dorigo, "Self-organised feedback in human swarm interaction," Tech. Rep. TR/IRIDIA/2012-016, IRIDIA, Université Libre de Bruxelles, Brussels, Belgium, December 2012.

[28] L. L. André Doyon, *Jacques Vaucanson: mécanicien de génie*. Presses Universitaires de France, 1967.

[29] I. Kato, S. Ohteru, H. Kobayashi, K. Shirai, and A. Uchiyama, *On Theory and Practice of Robots and Manipulators: Volume I*, ch. Information-Power Machine with Senses and Limbs, pp. 11–24. Vienna: Springer Vienna, 1974.

[30] J. Solis, K. Taniguchi, T. Ninomiya, T. Yamamoto, and A. Takanishi, "Development of waseda flutist robot WF-4RIV: implementation of auditory feedback system," in *2008 IEEE International Conference on Robotics and Automation, ICRA 2008, May 19-23, 2008, Pasadena, California, USA*, pp. 3654–3659, 2008.

[31] M. T. Takashima, S., "Control of an automatic performance robot of saxophone: performance control using standard MIDI files," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Workshop: Musical Performance and Its Applications*, 2006.

[32] J. Solis, T. Ninomiya, K. Petersen, M. Takeuchi, and A. Takanishi, "Development of the anthropomorphic saxophonist robot was-1: Mechanical design of the simulated organs and implementation of air pressure feedback control," *Advanced Robotics*, vol. 24, no. 5-6, pp. 629–650, 2010.

[33] S. H.-S. Y. A. M. S. M. Kuwabara, H., "The development of a violin musician robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Workshop: Musical Performance and Its Applications*, 2006.

[34] M. Kajitani, "Development of musician robots," *Journal of Robotics and Mechatronics*, vol. 1, pp. 254–255, 1989.

[35] K. Shibuya, S. Matsuda, and A. Takahara, "Toward developing a violin playing robot - bowing by anthropomorphic robot arm and sound analysis," in *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pp. 763–768, Aug 2007.

[36] G. Weinberg and S. Driscoll, "Toward robotic musicianship," *Comput. Music J.*, vol. 30, pp. 28–45, Dec. 2006.

[37] A. N. Kawamura, A. Alford, S. Northrup, K. Kawamura, K. w. Chan, and J. Barile, "Music playing robot," in *Proc. 1999 International Conference on Field and Service Robotics*, pp. 29–31, 1999.

[38] G. Weinberg, G. Hoffman, R. Nikolaidis, and R. Aim, "Shimon + zoozbeat: An improvising robot musician you can jam with," in *ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies: Adaptation*, SIGGRAPH ASIA '09, (New York, NY, USA), pp. 84–84, ACM, 2009.

[39] G. Hoffman and G. Weinberg, "Interactive improvisation with a robotic marimba player," *Autonomous Robots*, vol. 31, no. 2, pp. 133–153, 2011.

[40] M. Bretan, M. Cicconet, R. Nikolaidis, and G. Weinberg, "Developing and composing for a robotic musician using different modes of interaction," in *Non-Cochlear Sound: Proceedings of the 38th International Computer Music Conference, ICMC 2012, Ljubljana, Slovenia, September 9-14, 2012*, 2012.

[41] T. Mizumoto, R. Takeda, K. Yoshii, K. Komatani, T. Ogata, and H. G. Okuno, "A robot listens to music and counts its beats aloud by separating music from counting voice," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 1538–1543, Sept 2008.

[42] M. Bretan and G. Weinberg, "Chronicles of a robotic musical companion," in *14th International Conference on New Interfaces for Musical Expression, NIME 2014, London, United Kingdom, June 30 - July 4, 2014*, pp. 315–318, 2014.

[43] L. Maes, G. Raes, and T. Rogers, "The man and machine robot orchestra at logos," *Computer Music Journal*, vol. 35, pp. 28–48, Dec 2011.

[44] "Robot Band Backs Pat Metheny on Orchestrion Tour." `http://www.wired.com/2010/01/orchestrion/`. Date Accessed: 2016-02-23.

[45] Y. Uozumi, M. Takahashi, and R. Kobayashi, "A musical framework with swarming robots," in *Computer Music Modeling and Retrieval. Sense of Sounds* (R. Kronland-Martinet, S. Ystad, and K. Jensen, eds.), vol. 4969 of *Lecture Notes in Computer Science*, pp. 360–367, Springer Berlin Heidelberg, 2008.

[46] A. Albin, G. Weinberg, and M. Egerstedt, "Musical abstractions in distributed multi-robot systems," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 451–458, Oct 2012.

[47] T. Blackwell, "Swarm music: improvised music with multi-swarms," in *Proc. AISB Symposium in Artificial Intelligence and Creativity Arts Science*, pp. 41–49, 2003.

[48] J. McLurkin, “Stupid robot tricks: A behavior-based distributed algorithm library for programming swarms of robots,” Master’s thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 2004.

[49] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks*. Princeton series in applied mathematics, Princeton (N.J.): Princeton University Press, 2010.

[50] O. Krolyi, *Introducing music*. Middlesex, England: Penguin Books, first ed., 1965.

[51] U. Michels, *dtv Atlas zur Musik - Band I*. München: Deutcher Taschenbuch Verlag, first ed., 1977.

[52] B. Benward and M. Saker, *Music in Theory and Practice*. McGraw-Hill, first ed., 2003.

[53] W. Piston, *Harmony*. W. W. Norton and Company, fifth ed., 1987.

[54] J.-P. Rameau, *Traité de l’harmonie réduite à ses principes naturels*. Jean-Baptiste-Christophe Ballard.

[55] R. Olfati-Saber, “Near-identity diffeomorphisms and exponential epsi;-tracking and epsi;-stabilization of first-order nonholonomic se(2) vehicles,” in *American Control Conference, 2002. Proceedings of the 2002*, vol. 6, pp. 4690–4695 vol.6, 2002.

[56] S.-i. A. Meng Ji and M. Egerstedt, “Role-assignment in multi-agent coordination,” *International Journal of Assistive Robotics and Mechatronics*, vol. 7, pp. 32–40, March 2006.

[57] “K-Team Corporation - Mobile Robotics.” `http://www.k-team.com/`. Date Accessed: 2015-09-15.

[58] “OptiTrack - Motion Capture Systems.” `http://www.optitrack.com`. Date Accessed: 2015-09-15.

[59] “The Robotarium Project.” `www.robotarium.org`. Date Accessed: 2016-02-23.

[60] D. Pickem, M. Lee, and M. Egerstedt, “The GRITSBot in its Natural Habitat - A Multi-Robot Testbed,” in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pp. 4062–4067, May 2015.